

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:34:56 ; Search time 6.62013 Seconds
(without alignments)
4310.252 Million cell updates/sec

Title: US-09-926-799-6
Perfect score: 2896
Sequence: 1 MKMASNDAPSDGAGLVLP.....VNPFYTLAPMGTCNGRRRIQ 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1953.5	67.5	539	1	COAT LORDV
2	1123.5	38.8	546	1	COAT SOUV3
3	284.5	9.8	2344	1	POLN_RHDV
4	281.5	9.7	702	1	COAT SMSV1
5	275.5	9.5	671	1	COAT FCVF9
6	269	9.3	703	1	COAT SMSV4
7	262	9.0	668	1	COAT FCVC6
8	261.5	9.0	668	1	COAT FCVF4
9	255.5	8.8	2208	1	POLN_MANCV
10	140	4.8	2193	1	POLG CX16T
11	137.5	4.7	2194	1	POLG HE701
12	137	4.7	2193	1	POLG CX16G
13	128.5	4.4	2205	1	POLG_POL2W
14	126.5	4.4	2207	1	POLG_POL2L
15	125	4.3	2175	1	POLG BOVEV
16	121.5	4.2	2183	1	POLG CXB4E
17	119	4.1	2332	1	POLG FMDVO
18	118.5	4.1	2183	1	POLG_CXB4J
19	116.5	4.0	901	1	POLG_ENMG3
20	116.5	4.0	2333	1	POLG_FMDV1
21	116	4.0	1197	1	CAMI_BRARE
22	116	4.0	2193	1	POLG HE71M
23	114	3.9	2206	1	POLG_POL3L
24	112.5	3.9	932	1	CDGT_HUMAN
25	110.5	3.8	695	1	CICA_PAEPP
26	110.5	3.8	2201	1	POLG_CXA9
27	109.5	3.8	1324	1	SALI_HUMAN
28	109	3.8	2193	1	POLG HE71B
29	109	3.8	2206	1	POLG_POL1M
30	108.5	3.7	834	1	POLG_ENMGO
31	108.5	3.7	1011	1	POLG_FMDVT
32	108.5	3.7	2063	1	NC06_HUMAN
33	108.5	3.7	2332	1	POLG_FMDVA

34	108	3.7	2415	1	PGCA_HUMAN	P16112	homo sapien
35	107.5	3.7	487	1	DBD2_YEAST	P46957	saccharomyc
36	107	3.7	2208	1	POLH_POL1M	P03300	p genome po
37	106.5	3.7	2206	1	POLG_POL32	P06209	poliovirus
38	106	3.7	1159	1	N124_SCHPO	Q09904	schizosacch
39	105.5	3.6	435	1	CLPX_CLOPE	Q8xkk2	clostridium
40	105.5	3.6	2292	1	POLG_EMCVB	P17593	e genome po
41	104	3.6	621	1	AFS_MOUSE	Q9jid9	mus musculu
42	104	3.6	1794	1	YDC9_SCHPO	Q10172	schizosacch
43	103.5	3.6	2187	1	POLG_CXB2O	Q9vlg5	c genome po
44	102.5	3.5	1953	1	BNI1_YEAST	P41832	saccharomyc
45	102.5	3.5	2184	1	POLG_CXB6S	Q9ql88	c genome po

ALIGNMENTS

RESULT 1

COAT LORDV STANDARD; PRT; 539 AA.
AC P54635;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Coat protein (Capsid protein).
OS Lordsdale virus (Human enteric calicivirus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=82658;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=96005060; PubMed=7561776;
RA Diggle, K.E., Hambden, P., Camilleri, O., Clarke, I.N.,
Human enteric caliciviridae: the complete genome sequence and
expression of viral-like particles from a genetic group II small
round structured virus.
RL J. Gen. Virol. 76:2349-2355 (1995).
CC - SIMILARITY: TO CAPSID PROTEIN OF OTHER CALCIVIRUSES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

CC EMBL; X86557; CAA60255.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 531 531 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 539 AA; 58775 MW; 585C3E7F2C5FD21 CRC64;

Query Match 67.5%; Score 1953.5; DB 1; Length 539;

Best Local Similarity 66.2%; Pred. No. 6.6e-136;
Matches 365; Conservative 74; Mismatches 93; Indels 19; Gaps 4;

QY	1	MKMASNDAPSDGAGLVLP	INNEAMALDPVAGAAIAAPLTGQNIIDPWIMNPFVQAP	60
DB	1	MKMASNDANPDGSAANLVPEVNEVMALEPVVGAIAAPVAGQNVDPWRNNFVQAP	60	
QY	61	GGEFTVSPRNPGEVLLNLLELGPENPYLAHARMYNGYAGFEVQVVLGNAGFTAGKII	120	
DB	61	GGEFTVSPRNPAGEILWSAPLGLPOLNLYSLHSLRMVNGYAGFEVQVVLGNAGFTAGKVI	120	
QY	121	FAATPPNFPIDNLSAAOITMCPHVIVDVRQLEPVLNLPMPDVRNNFFHYNQSDSLRLIA	180	
DB	121	FAAVPPNFPTEGLSPFSQVTMPFHFIIVDVRQLEPVLNLPMPDVRNNFFHYNQANDSTKLIA	180	

```

RN      SEQUENCE FROM N.A.
RP      MEDLINE=93142023; PubMed=8380940;
RX      Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;
RT      "Sequence and genome organization of a human small round-structured
RT      (Norwalk-like) virus.";
RL      Science 259:516-519 (1993).
CC      -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; L07418; AAA92984.1; -.
CC      PIR; B37491; B37491.
CC      InterPro; IPR004005; Calici_coat.
CC      InterPro; IPR008975; Viral_cap_coat.
CC      Pfam; PF00915; Calici_coat; 1.
CC      Coat protein; Glycoprotein.
KW      CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ      SEQUENCE 546 AA; 58774 MW; C62B2A85AD4B05EA CRC64;

Query Match 38.8%; Score 1123.5; DB 1; Length 546;
Best local similarity 43.0%; Pred. No. 5.8e-75;

```


Db 508 QOTSRTLPASVG 519

RESULT 5

COAT FCVF9 STANDARD; PRT; 671 AA.

AC P27406;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 01-JUN-1994 (Rel. 29, Last annotation update)

DE Coat protein (Capsid protein).

GN C9P76.

OS Feline calicivirus (strain F9) (FCV).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC Vesiviruses.

OC NCBI_TaxID=11981;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92339673; PubMed=1633955;

RA Carter M.J., Milton I.D., Meanger J., Bennett M., Gaskell R.M.,

RA Turner P.C.;

RT "The complete nucleotide sequence of a feline calicivirus.";

RL Virology 190:443-448(1992).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=92339673; PubMed=1633955;

RA Meanger J., Carter M.J., Gaskell R.M., Turner P.C.;

RA "Cloning and sequence determination of the feline calicivirus strain F9.";

RT Biochem. Soc. Trans. 20:268-268(1992).

CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

CC EMBL; M86379; AAA79327.1; .

CC EMBL; Z11536; CAA7636.1; .

CC PIR; B43382; VCVWF9.

CC InterPro; IPR004005; Calici_coat.

CC InterPro; IPR008975; Viral_cap_coat.

CC Pfam; PF00915; Calici_coat; 1.

CC Coat protein; Glycoprotein.

CC CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC SEQUENCE 671 AA; 73441 MW; 33BEE86D8370D5ES CRC64;

CC

Query Match 9.5%; Score 275.5; DB 1; Length 671;

Best Local Similarity 22.3%; Pred. No. 1.4e-12;

Matches 123; Conservative 70; Mismatches 174; Indels 185; Gaps 23;

QY 67 SPRNPGVLLNLELGPINPYLAHRYNGYAGGFVQVVLGNAPTAGKIIPAAIPP 126

DB 179 STSETQKILFQSLGILLNPYLEHLAKLYVAGSIEVRFSGVFGGKLAIVVPP 238

QY 127 NFFPINLSAQITMCPHIVDQLEPNLMPDVNNFHHVQNSDRLRLIAMLYTPL 186

DB 239 G--VDFVQSTMLQPHVLFDAQVEPVIFCLPDLRSTLYLHMSDTR--TSLVIMVNDL 295

QY 187 RANNSGDDVFTVSC--RVLTTRSPDPSFNFLVPP-----TVES-----KT----- 224

DB 296 -INPYANDANSGCCIVTETKGPDKFHLKPPGSMLTGHSIPSDLPKTSLLMIGNRY 354

QY 225 -----KPTFL-----PI-LTISEMSNSRFPPIESLHTSP 253

DB 355 WSDITDFVIRPFVQANRHDFNQETAGWSTFRFRPISVTITEQNAKLGIGV-----A 408

QY 254 TENIVVQCONG--RVLTDLGELMGTTQLLPSSQICAFRGVLTSTSRASQADATATRLFN 311

DB 409 TDYIVFGIPDGWPDITIPGEL-----IPAGDYAITNGTGN----- 443

QY 312 YMHVQDLNLTGTPYPAEDIPGLTTPDRGKGVGVASQORNLDSTTRAHEAKVDTAGRP 371

DB 444 -----DITATGYDTADIK--NNTNFRGMVYCGSLQRAWG-----DKKISNTAFIT 488

QY 372 TPKLGSLSTSDSD-----PDNQPTKFTPVGIGVDNEAEFFQOWSLPDYSQGFTH 422

DB 489 TATL-----DGNNNKINPCNTIDQSKIVVFNQNHVGGKKAQTSDDTLALLGYTG----- 537

QY 423 NMLAPAVAFNPFGEQLL-----PFRSQLPSSGSRNGVLDCLVPQEWVH--FYQBS 473

DB 538 -----IGEQAIGSDRDRVVRISTLPETGARGG-----NHPIFYKNS 573

QY 474 AP-----AQOVALRVYNP-----DTGKVLFEAKLHLKLGPM 506

DB 574 IKLGVIRSIDVFNSQLHTSRQLSLNHLVLLPPDSPFAVYRIIDNSGWFIDIGDSGFSF 633

QY 507 IANNGDSPITVP 518

DB 634 VGVSGFGKLEFP 645

RESULT 6

COAT SMSV4 STANDARD; PRT; 703 AA.

AC P36285;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Coat protein (Capsid protein).

OS San Miguel sea lion virus (serotype 4) (SMSV 4).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC Vesiviruses.

OC NCBI_TaxID=36407;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92410750; PubMed=1529644;

RA Neill J.D.;

RT "Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lion virus: identification of conserved and non-conserved amino acid sequences among calicivirus capsid proteins.";

RL Virus Res. 24:211-222(1992).

CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

CC EMBL; M87482; AAA16220.1; .

CC PIR; C48562; C48562.

CC InterPro; IPR004005; Calici_coat.

CC InterPro; IPR008975; Viral_cap_coat.

CC Pfam; PF00915; Calici_coat; 1.

CC Coat protein; Glycoprotein.

CC CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC SEQUENCE 703 AA; 77721 MW; C5DAD823B261073 CRC64;

CC

Query Match 9.3%; Score 269; DB 1; Length 703;

Best Local Similarity 32.9%; Pred. No. 4.4e-12;

Matches 71; Conservative 38; Mismatches 89; Indels 18; Gaps 7;


```

1647 EMEGNSNPE-----PKQSNMNMVVDPP-----PGTGTPTTSHVVVANPEQNGAAQRL 1695
59 ---APGG-----EFTVSPRNSPEVILLNLELGEIPEINPYLAHLARMYNG 98
1696 LAVATGAIQSNVPEAIRNCFAVETFAWDRMPTGTFGLSISLHNINIPYSHLSGMAG 1755
99 YAGGFEVQVVLGNAFTAGKIIPAAIPNFFPIDNLSAAQITMCPHVIIVDVRQLEPNLPM 158
1756 WGSFVRLUSISGSGVFAIRIASVTPFG--VDPSSIRDPGLVPHAFVADARITEPVSEMI 1813
159 PDVRNFFHYNQSDSRSLRLIAMLYPLRANNSGDDVFTVSCRVLTRPSPDFSNFLVPP 218
1814 PDVRAVDYHRMDCGAETCSLGEFWVYQPLNPFSTTAVSTCWVSVEVKPGDGFCLLRPP 1873
219 TVESKYPFTLPILTITISEMSNSRFPVPIESLHSTPTENIVVOQNGRVTLDGLMGTOL 278
1874 G-----QWENG-VSPEG-----L 1886
279 LPSQICAFRG-----VLTRSTRASDQADTATPRLFN-----YYHHVOLDNLNGT 323
1887 LPRRLGYSRGNRVGGLVVGMLVAEHKQVN-----RHFNSNSVTFGW-----STA 1931
324 PYDP--AEDIPG-----PLGTPDRGKVGFGVASQORNLDSSTRAHEAKVDTTAGR 370
1932 PVNMAAEIVTNOAHSRHWLSIGAQN--KGFLFPGIPNHPFDPSCASTVVGAMDTSLGG 1990
371 FTKLGG---SLEISTSDDDQNOPT---KFTPV---GIGVDNEAEFQWMLP--- 414
1991 -RPSTGCGPAISQNNGDVYENDTFSVMEATYDPLTSGTGVALTNSINPASLALVRISN 2049
415 ---DYSG-----QFTHNNVLAPAVAP-----N 433
2050 NDETSFGFANDKNNVQMSWEMYGTQNGIRGQVTPMSGTNYTFTSTGANTLVLMQERMLS 2109
434 FPGQLLFFRQLSPSSGGR--SNGVLCLVQEWQVHEFYQESAPAQTOVALVRYVNPDTGK 492
2110 YDGHQAILYQLBERTAYFQNDIVN--IFENSWAVENVTNSASFQIG---IRPD--- 2160
493 VLFEAKLHKLGFMTIANNQDSPTVPPNGYPRFESWYVNPFTYLA PMG-----TGN-GRR 545
2161 -----GYM--VTGSGIGNVVPLEPETRFQ-----YVGILFLSALSGPSEGMGRA 2203
546 R 546
2204 R 2204

RESULT 10
POLG_CX16T
ID POLG_CX16T STANDARD; PRT; 2193 AA.
AC Q9QF31;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (PIA); Coat protein VP2 (PIB); Coat protein VP3 (PIC); Coat protein VP1 (PID); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VPG (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].
OS Coxsackievirus A16 (strain Taiwan/5079/98).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OC NCBI_TaxID=231417;
[1]
SEQUENCE FROM N.A.
MEDLINE=21427116; PubMed=11536241;
Yan J.-J., Su I.-J., Chen P.-F., Liu C.-C., Yu C.-K., Wang J.-R.;
"Complete genome analysis of enterovirus 71 isolated from an outbreak in Taiwan and rapid identification of enterovirus 71 and coxsackievirus A16 by RT-PCR."
J. Med. Virol. 65:331-339(2001).
CC RT
CC -!- FUNCTION: It is thought that the P2C protein attaches to vesicular membranes and is associated with viral RNA synthesis.

```

CC -|- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
 CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
 CC -|- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the
 CC poliovirus polypeptide. In other picornavirus reactions Glu may be
 CC substituted for Gln, and Ser or Thr for Gly.
 CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA) (N).
 CC -|- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AF17911; AAD55085.1; .
 CC HSP; P03300; IPOV.
 CC InterPro; IPR003593; AAA_ATPase
 CC InterPro; IPR004004; Calci_pol_hel.
 CC InterPro; IPR009003; Cys_Ser_trypsin.
 CC InterPro; IPR000199; Pept_3C_picorn.
 CC InterPro; IPR000081; Peptidase_C3.
 CC InterPro; IPR003138; Pico_P1A.
 CC InterPro; IPR002527; Pico_P2B.
 CC InterPro; IPR001676; Rhv.
 CC InterPro; IPR000605; RNA_helicase.
 CC InterPro; IPR007095; RNA_pol_DS_PS.
 CC InterPro; IPR001205; RNA_pol_P3D.
 CC InterPro; IPR007094; RNA_pol_P5vir.
 CC InterPro; IPR008975; Viral_cap_coat.
 CC Pfam; PF00548; Cys-protease_3C; 1.
 CC Pfam; PF02226; Pico_P1A; 1.
 CC Pfam; PF00947; Pico_P2A; 1.
 CC Pfam; PF01552; Pico_P2B; 1.
 CC Pfam; PF00073; rhv; 3.
 CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
 CC Pfam; PF00910; RNA_helicase; 1.
 CC PRINTS; PR00918; CALICVIRUSNS.
 CC ProDom; PD001125; Cys_protease_3C; 1.
 CC ProDom; PD001306; Pico_P2A; 1.
 CC ProDom; PD001274; Pico_P2B; 1.
 CC SMART; SM00382; AAA; 1.
 CC RNA-directed RNA polymerase; Core protein; Transferase;
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
 CC Lipoprotein.
 CC CHAIN 2 69 COAT PROTEIN VP4.
 CC CHAIN 70 323 COAT PROTEIN VP2.
 CC CHAIN 324 565 COAT PROTEIN VP3.
 CC CHAIN 566 862 COAT PROTEIN VP1.
 CC CHAIN 863 1012 CORE PROTEIN P2A.
 CC CHAIN 1013 1111 CORE PROTEIN P2B.
 CC CHAIN 1112 1440 CORE PROTEIN P2C.
 CC CHAIN 1441 1526 CORE PROTEIN P3A.
 CC CHAIN 1527 1548 GENOME-LINKED PROTEIN VPg.
 CC CHAIN 1549 1731 GENOME-DIRECTED RNA POLYMERASE.
 CC CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE.
 CC CHAIN 2 2 N-myristoyl glycine (in host) (By
 CC LIPID similarity).
 CC ACT_SITE 1695 1695 PROTEASE (POTENTIAL).
 CC ACT_SITE 1709 1709 PROTEASE (POTENTIAL).
 CC SEQUENCE 2193 AA; 243182 MW; 927839DB58F61E7F CRC64;
 CC Query Match 4.8%; Score 140; DB 1; Length 2193;
 CC Best Local Similarity 21.9%; Pred.No. 0.063;
 CC Matches 83; Conservative 41; Mismatches 133; Indels 122; Gaps 18;
 CC 57 VQAGGFTVSPRSPGVLNLELGPINPYLAHLARMYNGYAGGFEVQVVLGNAFTA 116

Db 398 VOSKTGELCAAFRADPGRD-----GPMQSTILGQLCRYTQWSGLSEVTFMFAGSFMT 451
 Qy 117 GKIIIPAAIPP--NFIIDNLSAAQITWCPHVIVDVVQLEPVLNPLMPDVRNPFPHYQGS 174
 Db 452 GKMLIATYPPCGNVPADRTAMLTG--HVIWDFGLQSSVTLVVPWISNT--HYR----- 501
 Qy 175 RLRLIAMLYTPLRANSGDVFTVSCRVLTRPSDFSNFLVP-----PT-----VESKTK 225
 Db 502 -----AHARAGYFDYTTGIIITW-----YQTNVYVPIGAPTTAYIVALAAQD 545
 Qy 226 PFTLPIITISEMSNSRFPVPIE---SLHTSPTEIVVQCGRVTLDGELMGTQLLP 282
 Db 546 NFMKLCCKDTE-----DIEQTANIQGDPIADMIDQTVNNQV---NRSLTALQVLP-- 592
 Qy 283 ICAPRGVLTSTRSRASDAQDTATPRLFNYYHVQLDNLNGTYPDPAEDIPGLGT---PD 339
 Db 593 -----TAADTEASSHR-----LGTGVVPA 611
 Qy 340 FRGKVFGV---ASQNLDSST--TRAHAKVDTTAGRTPKLGLSLSEISTDSDDFDQNP 394
 Db 612 LQAETGASSNASDKNLIETRCVLNHHSTQETAGNFFSRAGLVSIIT-----MPT- 662
 Qy 395 FFPVIGIGVDNEAEFOQWSL 413
 Db 663 -----TGQNTDGYVNWDI 676
 PRT; 2194 AA.
 POLG_HE701 STANDARD; PRT; 2194 AA.
 ID POLG_HE701
 AC P32537;
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein (Contains: Coat protein VP1 (P1A); Coat protein VP2
 DE (P1B); Coat protein VP3 (P1C); Coat protein VP4 (P1A); Core protein
 DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-
 DE linked protein VPg (P3B); Picornain 3C (SC 3.4.22.28) (Protease 3C)
 DE (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)).
 OS Human enterovirus 70 (strain J670/71) (EV 70).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Enterovirus.
 OC NCBI TaxID=31915;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91037960; PubMed=2172447;
 RA Ryan M.D., Jenkins O., Hughes P.J., Brown A., Knowles N.J., Booth D.,
 RA Minor P.D., Almond J.W.,
 RT "The complete nucleotide sequence of enterovirus type 70:
 RT relationships with other members of the picornaviridae.";
 RL J. Gen. Virol. 71:2291-2299(1990).
 CC -|- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
 CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
 CC -|- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the
 CC poliovirus polypeptide. In other picornavirus reactions Glu may be
 CC substituted for Gln, and Ser or Thr for Gly.
 CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA) (N).
 CC -|- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AF17911; AAD55085.1; .
 CC HSP; P03300; IPOV.
 CC InterPro; IPR003593; AAA_ATPase
 CC InterPro; IPR004004; Calci_pol_hel.
 CC InterPro; IPR009003; Cys_Ser_trypsin.
 CC InterPro; IPR000199; Pept_3C_picorn.
 CC InterPro; IPR000081; Peptidase_C3.
 CC InterPro; IPR003138; Pico_P1A.
 CC InterPro; IPR002527; Pico_P2B.
 CC InterPro; IPR001676; Rhv.
 CC InterPro; IPR000605; RNA_helicase.
 CC InterPro; IPR007095; RNA_pol_DS_PS.
 CC InterPro; IPR001205; RNA_pol_P3D.
 CC InterPro; IPR007094; RNA_pol_P5vir.
 CC InterPro; IPR008975; Viral_cap_coat.
 CC Pfam; PF00548; Cys-protease_3C; 1.
 CC Pfam; PF02226; Pico_P1A; 1.
 CC Pfam; PF00947; Pico_P2A; 1.
 CC Pfam; PF01552; Pico_P2B; 1.
 CC Pfam; PF00073; rhv; 3.
 CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
 CC Pfam; PF00910; RNA_helicase; 1.
 CC PRINTS; PR00918; CALICVIRUSNS.
 CC ProDom; PD001125; Cys_protease_3C; 1.
 CC ProDom; PD001306; Pico_P2A; 1.
 CC ProDom; PD001274; Pico_P2B; 1.
 CC SMART; SM00382; AAA; 1.
 CC RNA-directed RNA polymerase; Core protein; Transferase;
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
 CC Lipoprotein.
 CC CHAIN 2 69 COAT PROTEIN VP4.
 CC CHAIN 70 323 COAT PROTEIN VP2.
 CC CHAIN 324 565 COAT PROTEIN VP3.
 CC CHAIN 566 862 COAT PROTEIN VP1.
 CC CHAIN 863 1012 CORE PROTEIN P2A.
 CC CHAIN 1013 1111 CORE PROTEIN P2B.
 CC CHAIN 1112 1440 CORE PROTEIN P2C.
 CC CHAIN 1441 1526 CORE PROTEIN P3A.
 CC CHAIN 1527 1548 GENOME-LINKED PROTEIN VPg.
 CC CHAIN 1549 1731 GENOME-DIRECTED RNA POLYMERASE.
 CC CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE.
 CC CHAIN 2 2 N-myristoyl glycine (in host) (By
 CC LIPID similarity).
 CC ACT_SITE 1695 1695 PROTEASE (POTENTIAL).
 CC ACT_SITE 1709 1709 PROTEASE (POTENTIAL).
 CC SEQUENCE 2193 AA; 243182 MW; 927839DB58F61E7F CRC64;
 CC Query Match 4.8%; Score 140; DB 1; Length 2193;
 CC Best Local Similarity 21.9%; Pred.No. 0.063;
 CC Matches 83; Conservative 41; Mismatches 133; Indels 122; Gaps 18;
 CC 57 VQAGGFTVSPRSPGVLNLELGPINPYLAHLARMYNGYAGGFEVQVVLGNAFTA 116

DR ProDom; PD001125; Cys_protease_3C; 1.
 DR ProDom; PD001306; Pico_P2A; 1.
 DR ProDom; PD001274; Pico_P2B; 1.
 DR SMART; SM00382; AAA; 1.
 KW Polyprotein; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
 KW Lipoprotein. 2 69 COAT PROTEIN VP4.
 FT CHAIN 70 323 COAT PROTEIN VP2.
 FT CHAIN 324 565 COAT PROTEIN VP3.
 FT CHAIN 566 862 COAT PROTEIN VP1.
 FT CHAIN 863 1012 CORE PROTEIN P2A.
 FT CHAIN 1013 1111 CORE PROTEIN P2B.
 FT CHAIN 1112 1440 CORE PROTEIN P2C.
 FT CHAIN 1441 1526 CORE PROTEIN P3A.
 FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VP3.
 FT CHAIN 1549 1731 PICORNAIN 3C.
 FT CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE.
 FT LIPID 2 2 N-myristoyl glycine (in host) (By similarity).
 FT ACT SITE 1695 1695 PROTEASE (POTENTIAL).
 FT ACT SITE 1709 1709 PROTEASE (POTENTIAL).
 SQ SEQUENCE 2193 AA; 243209 MW; 0483BCE572A76E38 CRC64;

Query Match 4.7%; Score 137; DB 1; Length 2193;
 Best Local Similarity 20.6%; Pred. No. 0.1;
 Matches 92; Conservative 48; Mismatches 167; Indels 140; Gaps 20;

QY 57 VQAPGGEFTVSRNSPGEVLLNLELGPENPVLHARMYNGVAGFEVQVLAGNAFTA 116
 DB 398 VQSKTGELCAVFRADPG-----RNGPWQSTLGLCRYTQWQSGLEVTFFAGSFMAT 451
 QY 117 GKIIIFAAIAPP--NFTIDNLSSAQITMCPHVIVDVROLEPVNLPMPDVRNPFHYNQGS 174
 DB 452 GKMLIAYTPPGGVPAADRLTAMLTG---HVIWDFGLQSSVTLVPIWISNT--HYR----- 501
 QY 175 RURLIAMLYTPRANSGDDVTVCRLTRSPDFSNFLVPPVTSKTFPLTILTI 234
 DB 502 -----AHAKDGYDYTTTGTITW-----YQTYVVPV--IGAPTAYIVLALAA 543
 QY 235 SEMNSRFPVPVIESLHTSP-----ENIVQCGNRTLDGELMGTQLLSQICAFRG 288
 DB 544 QNFTWKLCKDTEDIEQSANIQDGIADMDQAVTSRV---GRALTSGLQVEP----- 592
 QY 289 VLTRSTRASDAQDTATPRLFNRYVHVDNLNLTGTPYDAEDIPGLPTG---PDRFGKVF 345
 DB 593 ---TAANTNASEHR-----LGTGLVPALQAAET 617
 QY 346 GVAS---QRNLDT--TRAHEAKVDTAGRTFKLGSLEISTSDDDFQONQTKFTPVGI 400
 DB 618 GASSNAQDENLITRCVLNHHSTQETTIGNFFSRAGLVSIIT-----MPT-----T 663
 QY 401 GVDNEAFQWSLPDYSQ-----FTHNNLAPAVAPNFPGEQLLFFRSQLPSSGG 451
 DB 664 GTQNTDGYNNWDI--DLMGYAQMRRKCELFY-----MRFDAETFFVAA 705
 QY 452 RSGVGLDCLVPOEWQHFQESAPACT 478
 DB 706 KPNGE---LVPQLLYMYVPPGAPKPT 729

RESULT 13

POLG_POL2W
 ID POLG_POL2W STANDARD; PRT; 2205 AA.
 AC P23069;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Picornain 2A (EC 3.4.22.29) (P2A); Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].

OS Poliovirus type 2 (strain W-2).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Enterovirus.
 OX NCBI_TaxID=12085;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90155230; PubMed=2154539;
 RA Pevear D.C., Oh C.K., Cunningham L.L., Calenoff M., Jubelt B.;
 RT "Localization of genomic regions specific for the attenuated, mouse-
 adapted poliovirus type 2 strain W-2.";
 RL J. Gen. Virol. 71:43-52(1990).
 CC FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
 Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
 CC CATALYTIC ACTIVITY: Selective cleavage of Gln-[Gly bond in the
 poliovirus polyprotein. In other picornavirus reactions Glu may be
 substituted for Gln, and Ser or Thr for Gly.
 CC CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 (RNA) (N).
 CC SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 each of which is composed of one copy each of proteins VP1, VP2,
 VP3, and VP4.
 CC PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; D00625; BAA00516.1; ALT_SEQ.
 PIR; A34032; GNNY2W.
 HSP; P03259; LPOV.
 MEROPS; C03.020; --
 InterPro; IPR003593; AAA ATPase.
 InterPro; IPR004004; Calici_pol_hel.
 InterPro; IPR009003; Cys Ser trypsin.
 InterPro; IPR001199; Pept_3C_picorn.
 InterPro; IPR000081; Peptidase_C3.
 InterPro; IPR003138; Pico_P1A.
 InterPro; IPR002527; Pico_P2B.
 InterPro; IPR001676; Rhv.
 InterPro; IPR006055; RNA helicase.
 InterPro; IPR007095; RNA_pol_DS_PS.
 InterPro; IPR001205; RNA_pol_P3D.
 InterPro; IPR007094; RNA_pol_Psvir.
 InterPro; IPR008975; Viral_cap_coat.
 Pfam; PF00548; Cys-protease-3C; 1.
 Pfam; PF02226; Pico_P1A; 1.
 Pfam; PF00947; Pico_P2A; 1.
 Pfam; PF01552; Pico_P2B; 1.
 Pfam; PF00073; rhv; 3.
 Pfam; PF00680; RNA dep RNA_pol; 1.
 Pfam; PF00910; RNA helicase; 1.
 PRINTS; PR00918; CALICIVIRUSN.
 ProDom; PD001125; Cys_protease_3C; 1.
 ProDom; PD001306; Pico_P2A; 1.
 ProDom; PD001274; Pico_P2B; 1.
 SMART; SM00382; AAA; 1.
 KW Polyprotein; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
 KW Lipoprotein. 2 69 COAT PROTEIN VP4.
 FT CHAIN 70 340 COAT PROTEIN VP2.
 FT CHAIN 341 578 COAT PROTEIN VP3.
 FT CHAIN 579 879 COAT PROTEIN VP1.
 FT CHAIN 880 1028 PROTEASE 2A.
 FT CHAIN 1029 1125 CORE PROTEIN 2B.
 FT CHAIN 1126 1454 CORE PROTEIN 2C.
 FT CHAIN 1455 1541 CORE PROTEIN 3A.

QY 76 -----LLNLELQPEINPIAH-----LARMNYAGGVEVQVVLNAGNAFTAGKIIF 121
 Db 413 NDAHSDTPILCISLSPASDPRLAHTMLGELIILNYTHWAGSLKFTFLFCGSMWATKLIIV 472
 QY 122 AALPPNFPIDNLSAQITWCPHVIVDVQLEPNLPMPPDVRRNFFH-----YNQGSDDR 175
 Db 473 SYAPPGAEPK-SRKEAMLGTHVIMDIGLOQSCWVPMWISNTTYRQTINDSFTEGGYIS 531
 QY 176 L-----RLIAMLYPLFRANNSGDDVFTVSCRVLTRPSPDFSNF----- 214
 Db 532 MFQTRVVVPLSTPRKWDILG---FVSACN-----DFSVLLRDTTHISQEAMPQGLG 581
 QY 215 -LVPPTVESKTPFTLPILITISEMSNRFPVPVIESLHTSPTENIVVQCGNVRVLDGELM 273
 Db 582 DLIEGVVEGVTRNALPTPLPANLPDQTSSGPAHSKET-PALTAV-----ETG 628
 QY 274 GTTOLPLSQAIFRGLV-----TRSTSRASD-----QADTATP-----RLFN--- 310
 Db 629 ANPLVPSTVQTRHVIOKRTSRSESVESFFARGACVAIEVDNDAPTGRASKLFSWKI 688
 QY 311 -YMHVQL-----DNLNGTPYD-----PAEDIPG--- 333
 Db 689 TYKDTVQLRKLEFFTYSRFDEFTFVTSNYTDANNGHALNQVQIMVIPGAPIGKW 748
 QY 334 -----PLGTPDFRGKV--EGVASQRNLDSSTRAHEAKVDTTAGRTFPKLG 376
 Db 749 NDYTWOTSSNPSPVFTYGAPPARISVPYVGIA---NAYSHFYDGFAGV-PLAGQASTEGD 804
 QY 377 SLEISTSDDF-----DQOQTKFT-----PVGIGVDNEAEFOQW-----SLP 414
 Db 805 SYLGASLNDGSLAVRVVNDNHPNTKLTIRVYMKPKHVRV-----WCPRPRAVP 856
 QY 415 DYSQGFTHMNLAPAVAPNFPGEQLLFFRSQLPSSGGRSNGVLDC---LVPOE----- 464
 Db 857 YVGPVGVYKDLAP-----LPKGLTYFGHQNKAVYTAGYKICNYHLATQEDLQNAV 911
 QY 465 --WVGHFYQESAPQVQVALRVYVNDPTGKVLFEAKLHLKLGFMWTIANNGSDPITVPPNGY 522
 Db 912 IMWIRDLVSKAQAQIDSIAR-CNCHTGVVYCESRRK---YYPVSFTGPTQYMEANBY 967
 QY 523 F--RPESWNPFTYPLAPMTGNGRRRIQ 548
 Db 968 YPARYQSHMLIGHGFASPGDCGGLRCQ 995

RESULT 15

POLG_BOVEV STANDARD; PRT; 2175 AA.
 AC P12915;
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DE Genome polypeptide [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VPG (P3B); Picornain 3C (SC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D (PC 2.7.7.48)].
 OS Bovine enterovirus (strain VG-5-27) (BEV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.
 OX NCBI_TaxID=12065;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88117392; PubMed=2828511;
 RA Earle J.A.P., Skuce R.A., Fleming C.S., Hoey E.M., Martin S.J.;
 RT "The complete nucleotide sequence of a bovine enterovirus.";
 RL J. Gen. Virol. 69:253-263 (1988).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 2-840.
 RX MEDLINE=95292108; PubMed=7773791;
 RA Smyth M., Tate J., Hoey E.M., Lyons C., Martin S.J., Stuart D.;
 RT "Implications for viral uncoating from the structure of bovine

enterovirus.";
 RL Nat. Struct. Biol. 2:224-231 (1995).
 CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
 CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-[Gly bond in the poliovirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + [RNA] (N).
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D00214; BAA24003.1; ALT_SEQ.
 DR FIR; A29824; GNNIBE.
 DR PDB; 1BEV; 16-SEP-98.
 DR MEROPS; C03.UBA; -.
 DR MEROPS; C03.UPB; -.
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR000199; Pept_3C_picorn.
 DR InterPro; IPR000081; Peptidase_C3.
 DR InterPro; IPR003138; Pico_P1A.
 DR InterPro; IPR002527; Pico_P2B.
 DR InterPro; IPR001676; Rbv.
 DR InterPro; IPR000605; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_Psvir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00548; Cys-protease-3C; 1.
 DR Pfam; PF02226; Pico_P1A; 1.
 DR Pfam; PF00947; Pico_P2A; 1.
 DR Pfam; PF01552; Pico_P2B; 1.
 DR Pfam; PF00073; rhv; 3.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICVIRUSNS.
 DR ProDom; PD001125; Cys_protease_3C; 1.
 DR ProDom; PD001306; Pico_P2A; 1.
 DR ProDom; PD001274; Pico_P2B; 1.
 DR Polyprotein; Coat protein; Core protein; Transferase; Thiol protease; Myristate; RNA-directed RNA polymerase; Hydrolase; Lipoprotein.
 KW 3D-structure; Lipoprotein.
 KW CHAIN 2 69
 FT CHAIN 70 317 COAT PROTEIN VP4.
 FT CHAIN 318 559 COAT PROTEIN VP2.
 FT CHAIN 560 840 COAT PROTEIN VP3.
 FT CHAIN 841 990 COAT PROTEIN VP1.
 FT CHAIN 991 1089 CORE PROTEIN P2A.
 FT CHAIN 1090 1419 CORE PROTEIN P2B.
 FT CHAIN 1420 1508 CORE PROTEIN P2C.
 FT CHAIN 1509 1531 CORE PROTEIN P3A.
 FT CHAIN 1532 1714 GENOME-LINKED PROTEIN VPG.
 FT CHAIN 1715 2175 PICORNAIN 3C.
 FT CHAIN 2 2 N-myriscyl glycine (in host).
 FT ACT SITE 1678 1678 PROTEASE (POTENTIAL).
 FT ACT SITE 1692 1692 PROTEASE (POTENTIAL).
 FT STRAND 26 26
 FT HELIX 36 38
 FT TURN 50 50
 FT HELIX 51 54
 FT STRAND 57 57

FT	STRAND	87	FT	STRAND	542	546
FT	TURN	88	FT	TURN	542	546
FT	STRAND	89	FT	TURN	550	551
FT	STRAND	90	FT	STRAND	574	574
FT	STRAND	94	FT	STRAND	578	578
FT	HELI	101	FT	STRAND	582	582
FT	TURN	102	FT	TURN	582	582
FT	HELI	103	FT	TURN	595	596
FT	TURN	105	FT	STRAND	597	598
FT	HELI	113	FT	HELI	600	602
FT	STRAND	123	FT	HELI	610	613
FT	STRAND	126	FT	HELI	619	619
FT	TURN	129	FT	STRAND	624	624
FT	STRAND	133	FT	STRAND	626	628
FT	STRAND	134	FT	HELI	629	629
FT	TURN	143	FT	HELI	630	634
FT	STRAND	147	FT	HELI	638	641
FT	TURN	151	FT	STRAND	644	644
FT	HELI	152	FT	TURN	647	649
FT	HELI	153	FT	STRAND	651	655
FT	TURN	155	FT	HELI	662	668
FT	TURN	156	FT	TURN	669	670
FT	HELI	157	FT	STRAND	671	685
FT	STRAND	167	FT	TURN	687	687
FT	TURN	180	FT	STRAND	689	690
FT	STRAND	185	FT	TURN	693	693
FT	TURN	186	FT	STRAND	699	705
FT	STRAND	187	FT	TURN	707	708
FT	TURN	197	FT	HELI	718	721
FT	STRAND	198	FT	STRAND	727	731
FT	TURN	209	FT	TURN	732	733
FT	HELI	212				
FT	HELI	215				
FT	STRAND	217				
FT	STRAND	219				
FT	TURN	220				
FT	HELI	224				
FT	TURN	227				
FT	STRAND	234				
FT	HELI	241				
FT	STRAND	245				
FT	TURN	246				
FT	STRAND	251				
FT	STRAND	256				
FT	TURN	265				
FT	STRAND	267				
FT	TURN	270				
FT	STRAND	273				
FT	STRAND	294				
FT	TURN	325				
FT	TURN	328				
FT	TURN	332				
FT	TURN	340				
FT	STRAND	343				
FT	TURN	344				
FT	STRAND	356				
FT	STRAND	359				
FT	TURN	360				
FT	HELI	361				
FT	TURN	364				
FT	TURN	365				
FT	STRAND	368				
FT	STRAND	375				
FT	TURN	376				
FT	TURN	377				
FT	TURN	381				
FT	HELI	382				
FT	STRAND	384				
FT	TURN	386				
FT	TURN	389				
FT	TURN	394				
FT	STRAND	397				
FT	TURN	407				
FT	HELI	410				
FT	HELI	415				
FT	TURN	420				
FT	TURN	421				
FT	STRAND	422				
FT	STRAND	423				
FT	TURN	430				
FT	TURN	440				
FT	TURN	441				
FT	STRAND	443				
FT	STRAND	445				
FT	HELI	461				
FT	TURN	465				
FT	TURN	466				
FT	STRAND	468				
FT	STRAND	473				
FT	STRAND	479				
FT	STRAND	493				
FT	TURN	499				
FT	TURN	502				
FT	HELI	503				
FT	STRAND	506				
FT	STRAND	510				
FT	TURN	515				
FT	STRAND	520				
FT	TURN	523				
FT	STRAND	528				
FT	TURN	537				
FT	TURN	539				
FT	TURN	540				

Query Match 4.3%; Score 125; DB 1; Length 2175;
Best Local Similarity 19.8%; Pred. No. 0.79; Mismatches 156; Indels 64; Gaps 10;
Matches 66; Conservative 47

QY 3 MASNDAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQQNIIDPWNNFVQAPGG 62
Db 313 IASNOGLETKFPG-----SYQFWTDEDCSPILPDFQFTLFIPIFGKVNLLLEIAQV 366
QY 63 EFTVSPRNSPGE-----VLINLEL-----GPEINPYLAHLARMYNGY 99
Db 367 ESILEANNREGVEGVYVIVSVQDALDAQIYALRLELGGSGPLSSLLGTLAKHYTQW 426
QY 100 AGGFEVQVVLGNAFTAGKIIFAAIPPNFIDNLSAAQITMCPHVIVDVRLQLEPVNLPMP 159
Db 427 SGSVEITCMFTGTFTMTGKVLAYTPGDMPR-NREAMLGTHVWVDFGLQSSITLVIP 485
QY 160 DVNNFFHYNQSDSRRLRLIAMLYTPLRANNSGDDVFTVSCRVLTRPSDFSFNLPVPT 219
Db 486 WISASHFR-----GVSNDVLYQYAAAGHVITWTQNMVIPP 524
QY 220 VESKTKPTLPLITISEMSNSRFPVPIESLHTSPTEINLVQCONGRV---TLDGELMGTT 276
Db 525 F-----PNTAGIIMMI-AAQPNFSFRIQKREDMTQTALQNDPGKMLKDAIDKOVAGA- 577
QY 277 QLLPSQICAFRCVLTSTST--RASDAQDTATPR 307
Db 578 -LVAGTTTSTHSVATDSTPALQAETGATSTAR 609

Search completed: June 1, 2004, 13:47:18
Job time : 8.62013 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:35:27 ; Search time 31.9054 Seconds

(without alignments)
5419.275 Million cell updates/sec

Title: US-09-926-799-6

Perfect score: 2896

Sequence: 1 MKMASNDAPSDGAGLVP.....VNPFYTLAPMGNGRRRIQ 548

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2855	98.6	548	12 Q8BC87	Q8bc87 human calic
2	2851	98.4	548	12 Q8PYA7	Q8pya7 human calic
3	2850	98.4	548	12 Q916B5	Q916b5 human calic
4	2836	97.9	548	12 Q8BC90	Q8bc90 human calic
5	2831	97.8	548	12 Q915D2	Q915d2 human calic
6	2822	97.4	548	12 Q66296	Q66296 calicivirus
7	2819	97.3	548	12 Q915D3	Q915d3 human calic
8	2815	97.2	548	12 Q8V0P5	Q8v0p5 human calic
9	2812	97.1	548	12 Q918A1	Q918a1 normalik-lik
10	2810	97.0	548	12 Q8V0P4	Q8v0p4 human calic
11	2802	96.8	548	12 Q8V768	Q8v768 normalik vir
12	2799	96.7	548	12 Q88291	Q88291 small round
13	2794	96.5	548	12 Q83880	Q83880 normalik vir
14	2786	96.2	548	12 Q917Z5	Q917z5 normalik-lik
15	2781	96.0	548	12 Q917Z2	Q917z2 normalik-lik
16	2780	96.0	548	12 Q91V40	Q91v40 human calic

17	2777	95.9	548	12 Q8V771	Q8v771 normalik vir
18	2776	95.9	548	12 Q917Z8	Q917z8 normalik-lik
19	2772	95.7	548	12 Q68291	Q68291 human calic
20	2770	95.6	548	12 Q918A4	Q918a4 normalik-lik
21	2768	95.6	534	12 Q9QAX4	Q9qax4 normalik-lik
22	2742	94.7	548	12 Q91V46	Q91v46 human calic
23	2727	94.2	548	12 Q80FK9	Q80fk9 human calic
24	2720	93.9	548	12 Q80FK8	Q80fk8 human calic
25	2720	93.9	548	12 Q80FK5	Q80fk5 human calic
26	2718	93.9	548	12 Q80FK6	Q80fk6 human calic
27	2715	93.8	548	12 Q80FL1	Q80fl1 human calic
28	2707	93.5	548	12 Q80FK3	Q80fk3 human calic
29	2704	93.4	548	12 Q80FK7	Q80fk7 human calic
30	2701	93.3	548	12 Q80FK4	Q80fk4 human calic
31	2695	93.1	548	12 Q80FK2	Q80fk2 human calic
32	2688	92.8	548	12 Q80FL2	Q80fl2 human calic
33	2688	92.8	548	12 Q80FK1	Q80fk1 human calic
34	2676	92.4	548	12 Q80FL0	Q80fl0 human calic
35	2173.5	75.1	547	12 Q918A7	Q918a7 normalik-lik
36	2171.5	75.0	547	12 Q8V775	Q8v775 normalik vir
37	2161.5	74.6	547	12 Q8V773	Q8v773 normalik vir
38	2148.5	74.2	535	12 Q917X4	Q917x4 normalik-lik
39	2147.5	74.2	535	12 Q916E6	Q916e6 human calic
40	2147.5	74.2	535	12 Q916E4	Q916e4 human calic
41	2145.5	74.1	535	12 Q917Y3	Q917y3 normalik-lik
42	2140.5	73.9	535	12 Q68104	Q68104 hawaii cali
43	2140.5	73.9	535	12 Q917Y9	Q917y9 normalik-lik
44	2139	73.9	550	12 Q918B3	Q918b3 normalik-lik
45	2137.5	73.8	535	12 Q917Y0	Q917y0 normalik-lik

ALIGNMENTS

RESULT 1

Q8BC87 PRELIMINARY: PRT; 548 AA.
 AC Q8BC87;
 DT 01-MAR-2003 (TremBLrel. 23, Created)
 DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
 DE Capsid protein.
 OS Human calicivirus NLV/Oberhausen 455/01/DE.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 OC NCBI TaxID=173921;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Hu/NLV/Oberhausen 455/01/DE;
 RA Kuenkel U., Schreier E.;
 RT "Characterization of recombinant human calicivirus circulating in Germany."
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF539440; AAN16373.1; ...
 DR InterPro; IPR004005; Calic_coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calic1_coat; 1.
 SQ SEQUENCE 548 AA; 59985 MW; 9797ED4783790887 CRC64;

Query Match 98.6%; Score 2855; DB 12; Length 548;
 Best Local Similarity 97.8%; Pred. No. 9.3e-218;
 Matches 536; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY	1	MKMASNDAPSDGAGLVP	PEINNEAMALDPVAGAAIAAPLTGQNIIDPWI	MNNFVQAP 60
Db	1	MKMASNDATPNDGAGLVP	PEINNEAMALDPVAGAAIAAPLTGQNIIDPWI	MNNFVQAP 60.
QY	61	GGFTVSPRSPGEVLLNLELGP	EINPYLAHLARMYNGYAGGFVQVVL	AGNAFTAGKII 120
Db	61	GGFTVSPRSPGEVLLNLELGP	EINPYLAHLARMYNGYAGGFVQVVL	AGNAFTAGKII 120
QY	121	FAAPPNPPIINLSAQITMCPH	IVDVROLEPNLMPDVNNPFHNNQGS	DSRLRIA 180

Db 121 FAAPFPIDNLSAAQITMCPHVIVDVRLQEPVNLPMFVDRNNFFHYNQSGDSRLRLIA 180
QY 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPVTVESKTKPFTLPILITISEMSNS 240
Db 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPVTVESKTKPFTLPILITISEMSNS 240
QY 241 RFPVPISLHTSPTENIVVQCGRVTLDGELMGTTQLLPSQICAFRGVLTRSTRASDQ 300
Db 241 RFPVPISLHTSPTENIVVQCGRVTLDGELMGTTQLLPSQICAFRGVLTRSTRASDQ 300
QY 301 ADTATPRLFNYYWHVQDNLNGTPYDPAEDIPGLGTPDFRGKVFQVASORNLDSITTRAH 360
Db 301 ADTATPRLFNYYWHVQDNLNGTPYDPAEDIPGLGTPDFRGKVFQVASORNLDSITTRAH 360
QY 361 EAKVDTTACGFTPKLGSLEISTDSDDFQDQNTKFTTPVGIGVDNEAFQWLSLDPYSGOF 420
Db 361 EAKVDTTACGFTPKLGSLEISTDSDDFQDQNTKFTTPVGIGVDNEAFQWLSLDPYSGOF 420
QY 421 THNNMLAPAVAPNFPGEQLLFFRSQLPSSGGRSNGVLDCLVPOEWVQHFYQESAPAQTOV 480
Db 421 THNNMLAPAVAPNFPGEQLLFFRSQLPSSGGRSNGVLDCLVPOEWVQHFYQESAPAQTOV 480
QY 481 ALVRYNPDGTGKVLFEAKLHKLGFMTIANGDSPIITVPPNGYFRFESWNPFFYTLAPMGT 540
Db 481 ALVRYNPDGTGKVLFEAKLHKLGFMTIANGDSPIITVPPNGYFRFESWNPFFYTLAPMGT 540
QY 541 GNGRRRIQ 548
Db 541 GNGRRRIQ 548

RESULT 2
Q9PYA7 PRELIMINARY; PRT; 548 AA.
ID Q9PYA7
AC Q9PYA7
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus strain Arg320.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OC NCBI_TaxID=106586;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Arg320;
RX MEDLINE=20129490; PubMed=10664391;
RA Jiang X., Espul C., Zhong W.M., Cuello H., Matson D.O.;
RT "Characterization of a novel human calicivirus that may be a naturally occurring recombinant";
RL Arch. Virol. 144:2377-2387(1999);
DR EMBL; AF190817; AAF13920.1; -;
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 53872 MW; 46DCC4042967CD79 CRC64;

Query Match 98.4%; Score 2851; DB 12; Length 548;
Best Local Similarity 98.2%; Pred. No. 1.9e-217;
Matches 538; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
QY 1 MKMASNDAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQNIIDPWIMNFFVOAP 60
Db 1 MKMASNDAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQNIIDPWIMNFFVOAP 60
QY 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHARMYNGYAGGFVQVVLGNATAGKII 120
Db 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHARMYNGYAGGFVQVVLGNATAGKII 120
QY 121 FAAPFPIDNLSAAQITMCPHVIVDVRLQEPVNLPMFVDRNNFFHYNQSGDSRLRLIA 180
Db 121 FAAPFPIDNLSAAQITMCPHVIVDVRLQEPVNLPMFVDRNNFFHYNQSGDSRLRLIA 180

QY 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPVTVESKTKPFTLPILITISEMSNS 240
Db 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPVTVESKTKPFTLPILITISEMSNS 240
QY 241 RFPVPISLHTSPTENIVVQCGRVTLDGELMGTTQLLPSQICAFRGVLTRSTRASDQ 300
Db 241 RFPVPISLHTSPTENIVVQCGRVTLDGELMGTTQLLPSQICAFRGVLTRSTRASDQ 300
QY 301 ADTATPRLFNYYWHVQDNLNGTPYDPAEDIPGLGTPDFRGKVFQVASORNLDSITTRAH 360
Db 301 ADTATPRLFNYYWHVQDNLNGTPYDPAEDIPGLGTPDFRGKVFQVASORNLDSITTRAH 360
QY 361 EAKVDTTACGFTPKLGSLEISTDSDDFQDQNTKFTTPVGIGVDNEAFQWLSLDPYSGOF 420
Db 361 EAKVDTTACGFTPKLGSLEISTDSDDFQDQNTKFTTPVGIGVDNEAFQWLSLDPYSGOF 420
QY 421 THNNMLAPAVAPNFPGEQLLFFRSQLPSSGGRSNGVLDCLVPOEWVQHFYQESAPAQTOV 480
Db 421 THNNMLAPAVAPNFPGEQLLFFRSQLPSSGGRSNGVLDCLVPOEWVQHFYQESAPAQTOV 480
QY 481 ALVRYNPDGTGKVLFEAKLHKLGFMTIANGDSPIITVPPNGYFRFESWNPFFYTLAPMGT 540
Db 481 ALVRYNPDGTGKVLFEAKLHKLGFMTIANGDSPIITVPPNGYFRFESWNPFFYTLAPMGT 540
QY 541 GNGRRRIQ 548
Db 541 GNGRRRIQ 548

RESULT 3
Q916E5 PRELIMINARY; PRT; 548 AA.
ID Q916E5
AC Q916E5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/Oberhausen 455/01/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OC NCBI_TaxID=173921;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/Oberhausen 455/01/DE;
RA Kuenkel U., Hoehne M., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF425768; AAL18861.1; -;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 55958 MW; 8BA72E4FDF0330D0 CRC64;

Query Match 98.4%; Score 2850; DB 12; Length 548;
Best Local Similarity 97.6%; Pred. No. 2.3e-217;
Matches 535; Conservative 10; Mismatches 3; Indels 0; Gaps 0;
QY 1 MKMASNDAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQNIIDPWIMNFFVOAP 60
Db 1 MKMASNDAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQNIIDPWIMNFFVOAP 60
QY 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHARMYNGYAGGFVQVVLGNATAGKII 120
Db 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHARMYNGYAGGFVQVVLGNATAGKII 120
QY 121 FAAPFPIDNLSAAQITMCPHVIVDVRLQEPVNLPMFVDRNNFFHYNQSGDSRLRLIA 180
Db 121 FAAPFPIDNLSAAQITMCPHVIVDVRLQEPVNLPMFVDRNNFFHYNQSGDSRLRLIA 180
QY 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPVTVESKTKPFTLPILITISEMSNS 240
Db 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPVTVESKTKPFTLPILITISEMSNS 240

```

QY 241 RPPVPIESLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPQICAFRGVLTSTSRASDQ 300
DB 241 RPPVPIESLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPQICAFRGVLTSTSRASDQ 300
QY 301 ADTATPRLFNYYWHVQVOLDNLNGTTPYDPAEDIPGLGTDPFRGKVFQVASQRNLDSTTRAH 360
DB 301 ADTATPRLFNYYWHVQVOLDNLNGTTPYDPAEDIPGLGTDPFRGKVFQVASQRNLDSTTRAH 360
QY 361 EAKVDTTACGRTPTKLGSLISTSDDDFQDQNTKFTPTVGVIGVDNEAEFOQWLSLPYSGOF 420
DB 361 EAKVDTTACGRTPTKLGSLISTSDDDFQDQNTKFTPTVGVIGVDNEAEFOQWLSLPYSGOF 420
QY 421 THNNMLAPAVANFPFGQQLLFRSOLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQTOV 480
DB 421 THNNMLAPAVANFPFGQQLLFRSOLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQTOV 480
QY 481 ALRVYVNPDTGKVLPEAKLHKGFTMTIANGDSPITVPNGYFRFESWVNPFTYLA PMGT 540
DB 481 ALRVYVNPDTGKVLPEAKLHKGFTMTIANGDSPITVPNGYFRFESWVNPFTYLA PMGT 540
QY 541 GNGRRRIQ 548
DB 541 GNGRRRIQ 548

```

RESULT 4

```

Q8BC90 PRELIMINARY; PRT; 548 AA.
AC Q8BC90;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/Herzberg 385/01/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OC NCBI_TaxID=207828;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Herzberg 385/01/DE;
RA Kuenkel U., Schreier E.;
RT "Characterization of recombinant human calicivirus circulating in Germany.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF539439; AAN16370.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 59954 MW; D0C02A8E89D6EBD6 CRC64;

```

```

Query Match 97.9%; Score 2836; DB 12; Length 548;
Best Local Similarity 97.3%; Pred. No. 3e-216; Mismatches 10; Indels 0; Gaps 0;
Matches 533; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

QY 1 MKWASNDAAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQQNIIDPWIMNPFVQAP 60
DB 1 MKWASNDAAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQQNIIDPWIMNPFVQAP 60
QY 61 GGEFTVSPRNSPGEVLLNLELGPINPYLAHARMYNGVAGFEVQVVLGNAGNAGTAKGII 120
DB 61 GGEFTVSPRNSPGEVLLNLELGPINPYLAHARMYNGVAGFEVQVVLGNAGNAGTAKGII 120
QY 121 FFAIIPNPFIDNLAAQITMCPHVIIVDVRQLEPVLNMPDVRNPNFHYNQGSLSRLRLIA 180
DB 121 FFAIIPNPFIDNLAAQITMCPHVIIVDVRQLEPVLNMPDVRNPNFHYNQGSLSRLRLIA 180
QY 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPTVESKTPFTLPILTISEMSNS 240
DB 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPTVESKTPFTLPILTISEMSNS 240
QY 241 RPPVPIESLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPQICAFRGVLTSTSRASDQ 300
DB 241 RPPVPIESLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPQICAFRGVLTSTSRASDQ 300

```

```

DB 241 RPPVPIESLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPQICAFRGVLTSTSRASDQ 300
QY 301 ADTATPRLFNYYWHVQVOLDNLNGTTPYDPAEDIPGLGTDPFRGKVFQVASQRNLDSTTRAH 360
DB 301 ADTATPRLFNYYWHVQVOLDNLNGTTPYDPAEDIPGLGTDPFRGKVFQVASQRNLDSTTRAH 360
QY 361 EAKVDTTACGRTPTKLGSLISTSDDDFQDQNTKFTPTVGVIGVDNEAEFOQWLSLPYSGOF 420
DB 361 EAKVDTTACGRTPTKLGSLISTSDDDFQDQNTKFTPTVGVIGVDNEAEFOQWLSLPYSGOF 420
QY 421 THNNMLAPAVANFPFGQQLLFRSOLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQTOV 480
DB 421 THNNMLAPAVANFPFGQQLLFRSOLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQTOV 480
QY 481 ALRVYVNPDTGKVLPEAKLHKGFTMTIANGDSPITVPNGYFRFESWVNPFTYLA PMGT 540
DB 481 ALRVYVNPDTGKVLPEAKLHKGFTMTIANGDSPITVPNGYFRFESWVNPFTYLA PMGT 540
QY 541 GNGRRRIQ 548
DB 541 GNGRRRIQ 548

```

RESULT 5

```

Q915D2 PRELIMINARY; PRT; 548 AA.
AC Q915D2;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Bitburg/289/01/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OC NCBI_TaxID=173924;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/Bitburg/289/01/DE;
RA Kuenkel U., Hoehne M., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF427112; AAL18864.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 59823 MW; A83BC58F87F93088 CRC64;

```

```

Query Match 97.8%; Score 2831; DB 12; Length 548;
Best Local Similarity 97.4%; Pred. No. 7.5e-216; Mismatches 7; Indels 0; Gaps 0;
Matches 534; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 1 MKWASNDAAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQQNIIDPWIMNPFVQAP 60
DB 1 MKWASNDAAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQQNIIDPWIMNPFVQAP 60
QY 61 GGEFTVSPRNSPGEVLLNLELGPINPYLAHARMYNGVAGFEVQVVLGNAGNAGTAKGII 120
DB 61 GGEFTVSPRNSPGEVLLNLELGPINPYLAHARMYNGVAGFEVQVVLGNAGNAGTAKGII 120
QY 121 FFAIIPNPFIDNLAAQITMCPHVIIVDVRQLEPVLNMPDVRNPNFHYNQGSLSRLRLIA 180
DB 121 FFAIIPNPFIDNLAAQITMCPHVIIVDVRQLEPVLNMPDVRNPNFHYNQGSLSRLRLIA 180
QY 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPTVESKTPFTLPILTISEMSNS 240
DB 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPTVESKTPFTLPILTISEMSNS 240
QY 241 RPPVPIESLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPQICAFRGVLTSTSRASDQ 300
DB 241 RPPVPIESLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPQICAFRGVLTSTSRASDQ 300
QY 301 ADTATPRLFNYYWHVQVOLDNLNGTTPYDPAEDIPGLGTDPFRGKVFQVASQRNLDSTTRAH 360

```

Qy	361	BAKVDTAGRPTPKLGSLEISTSDDFDQONQTKPTVCGIGVDNEAEFOQWSLPDYSGQF	420
Db	361	EAKVDTTSGRFTPKLGSLEITTESDDDFDQNPQTKPTFPVGVGVNDNEAEFOQWSLPNYSGQF	420
Qy	421	THNNMLAPAVAPNFPGEQLLFFRSQLPSSSGRSNGVLDCLVPQEWVQHIFYQSAPAQTV	480
Db	421	THNNMLAPAVAPNFPGEQLLFFRSQLPSSSGRSNGVLDCLVPQEWVQHIFYQSAPAQTV	480
Qy	481	ALVRYNPDGTGKVLFPKAKLHKLGFMTIANNQSDSPITVPNGYFRFESWNPFPVTLAPMGT	540
Db	481	ALVRYNPDGTGKVLFPKAKLHKLGFMTIANNQSDSPITVPNGYFRFESWNPFPVTLAPMGT	540
Qy	541	GNGRRRRIQ 548	
Db	541	GNGRRRRIQ 548	
RESULT 7			
Q915D3 PRELIMINARY; PRT; 548 AA.			
ID	Q915D3		
AC	Q915D3		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Capsid protein.		
OS	Human calicivirus HU/NLW/Berlin/226/01/DE.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;		
OC	Norovirus.		
OX	NCBI_TaxID=173923;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HU/NLW/Berlin/226/01/DE;		
PA	Kuenkel U., Hoehne M., Schreier E.;		
RT	"Molecular epidemiology of outbreaks of gastroenteritis associated		
RT	with Norwalk-like viruses in Germany.;"		
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF427111; AAL18863.1; -		
DR	InterPro; IPR004005; Calici_coat.		
DR	InterPro; IPR008975; Viral_Gap_coat.		
DR	Pfam; PF00915; Calici_coat; 1.		
DR	SEQUENCE 548 AA; 53844 MW; 391B68558ABDFE85 CRC64;		
Query Match 97.3%; Score 2819; DB 12; Length 548;			
Best Local Similarity 96.9%; Pred. No. 6.7e-215;			
Matches 531; Conservative 8; Mismatches 9; Indels 0; Gaps 0;			
Qy	1	MKMASNDAPNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQONIIDPWIMNFWQAP	60
Db	1	MKMASNDATPNDGAGLVPEINNEAMALDPVAGAAIAAPLTGQONIIDPWIMNFWQAP	60
Qy	61	GGEFTVSPRNSPGEVLLNLEIGPEINPYLAHARMYNGYAGGFVQVLAGNAFTAGKII	120
Db	61	GGEFTVSPRNSPGEVLLNLEIGPEINPYLAHARMYNGYAGGFVQVLAGNAFTAGKII	120
Qy	121	FAAIPNPEIDNLAAQITMCQPHVLDVQLEPNLPMDPVRNNEFHYNQGSDSRLRLIA	180
Db	121	FAAXPNPEIDNLAAQITMCQPHVLDVQLEPNLPMDPVRNNEFHYNQGSDSRLRLIA	180
Qy	181	MLYTPLRANNSGDDVFTVSCRVLTRPSDFGFNFVLPPTVESKTPFTLPLITISEMSNS	240
Db	181	MLYTPLRANNSGDDVFTVSCRVLTRPSDFGFNFVLPPTVESKTPFTLPLITISEMSNS	240
Qy	241	RFPVPIESLHSTPTENIVVQCONGAVTLDELGMGTQLLPSQICAFRGVLTSTSRASQD	300
Db	241	RFPVPIESLHSTPTENIVVQCONGAVTLDELGMGTQLLPSQICAFRGVLTSTSRASQD	300
Qy	301	ADTATPRLEPNYVHVDLNLNCTPDPAEDIDPGILGTPD FRGKVFGEVASQRNLDSTTRAH	360
Db	301	ADTATPRLEPNYVHVDLNLNCTPDPAEDIDPGILGTPD FRGKVFGEVASQRNLDSTTRAH	360
Qy	361	EAKVDTAGRPTPKLGSLEISTSDDFDQONQTKPTVCGIGVDNEAEFOQWSLPDYSGQF	420
Db	361	EAKVDTAGRPTPKLGSLEISTSDDFDQONQTKPTVCGIGVDNEAEFOQWSLPDYSGQF	420

Db	301	ADTATPLRFLNYWHIQLDNLNGTTPCDPAEDIPGLGTDPFRGKVGFGVASQRNPDSTTGAH	360
Qy	361	EAKVDTTAGRTPKLGSLEISTSDPFDQONQTKFTPVGIGVDNEAFQOWSLPDYSGQF	420
Db	361	EAKVDTTAGRTPKLGSLEISTSGDFDQONQTKFTPVGIGVDHEADFQOWSLPDYSGQF	420
Qy	421	THNNNLAPAVAPNPPGQLLFFRSQLPSSGGRSNGVLDCLPQEWVQHFYQESAPAQTV	480
Db	421	THNNNLAPAVAPNPPGQLLFFRSQLPSSGGRSNGILDCLPQEWVQHFYQESAPAQTV	480
Qy	481	ALRVYVNPDTGKVLFEAKLHKLGFMTIANGDSPITVPNGVFRPESWNNPFTYLA PMGT	540
Db	481	ALRVYVNPDTGRVLFEAKLHKLGFMTIANGDSPITVPNGVFRPESWNNPFTYLA PMGT	540
Qy	541	GNRRRIQ 548	
Db	541	GNRRRIQ 548	
RESULT 6			
ID	Q66296	PRELIMINARY; PRT; 548 AA.	
AC	Q66296;		
DT	01-NOV-1996 (TREMBLrel. 01, Created)		
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)		
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)		
DE	Capaid protein.		
OS	Calciocivirus.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae.		
NCBI_Taxid=	11975;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RP	STRAIN=TV24;		
RX	MEDLINE=94202329; PubMed=8151799;		
RT	Lew J.F.; Petric M. Kapikian A. Z., Jiang X., Estes M.K., Green K.Y.;		
RT	"Identification of Minireovirus as a Norwalk-like virus in pediatric		
RT	patients with gastroenteritis.";		
RL	J. Virol. 68:3391-3396(1994).		
DR	EMBL; U02030; AAA18930.1; -.		
DR	InterPro; IPR004005; Calici coat.		
DR	InterPro; IPR008975; Viral Cap coat.		
DR	Fram; PR00915; Calici coat_1.		
SQ	SEQUENCE 548 AA; 60034 MW; CCEE4A180E85008B CRC64;		
Query Match 97.4%; Score 2822; DB 12; Length 548;			
Best Local Similarity 96.7%; Pred. No. 3.9e-215;			
Matches 530; Conservative 11; Mismatches 7; Indels 0; Gaps 0;			
Qy	1	MKVASNDAPSNDGAGLVPEINNEAMALDPVAGAAIAAPLTGQNIIDPWIMNFWQAP	60
Db	1	MKVASNDAPSNDGAGLVPEINNEAMALEPVAGSAIAAPLTGQNIIDPWIMNFWQAP	60
Qy	61	GGFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGGFEVQVVLAGNAFTAGKII	120
Db	61	GGFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGGFEVQVVLAGNAFTAGKII	120
Qy	121	FAAIPNFPIDNLAAQITMCCHVIVDVROLEPNLMPDVRNPFHYNGSDSRRLRLTA	180
Db	121	FAAIPNFPIDNLAAQITMCCHVIVDVROLEPNLMPDVRNPFHYNGSDSRRLRLTA	180
Qy	181	MLYTLPRANNSGDDVFTVSCRVLTSPDFGFNLFVPPTVESKTKPFTLPILITSEMNS	240
Db	181	MLYTLPRANNSGDDVFTVSCRVLTSPDFGFNLFVPPTVESKTKPFTLPILITSEMNS	240
Qy	241	RFPVPVIESLHSTPENIVVQCGNRVTLDELMGTTLQLPSQICAFRGVITRSTSRASQ	300
Db	241	RFPVPVDSLHSTPENIVVQCGNRVTLDELMGTTLQLPSQICAFRGVITRSTSRASQ	300
Qy	301	ADTATPLRFLNYWHIQLDNLNGTTPCDPAEDIPGLGTDPFRGKVGFGVASQRNPDSTTGAH	360
Db	301	ADTPTPLRFLNYWHIQLDNLNGTTPCDPAEDIPGLGTDPFRGKVGFGVASQRNPDSTTGAH	360

QY 421 THNNMLAPAVANFPGEQLLFRSOLPSSGGRSNGVLDCLVPOEWWOHPYQESAPAQTOV 480
 DB 421 THNNMLAPAVANFPGEQLLFRSOLPSSGGRSNGVLDCLVPOEWWOHPYQESAPAQTOV 480
 QY 481 ALVRVYNPDGTGVLFEAKLHKLGFMTIANNGDSPITVPNGYFRFESWNPVYTLAPMGT 540
 DB 481 ALVRVYNPDGTGVLFEAKLHKLGFMTIANNGDSPITVPNGYFRFESWNPVYTLAPMGT 540
 QY 541 GNGRRRIQ 548
 DB 541 GNGRRRIQ 548

RESULT 8
 Q8VOP5 PRELIMINARY; PRT; 548 AA.
 AC Q8VOP5;
 DT 01-MAR-2002 (TREMELrel. 20, Created)
 DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Capsid protein.
 OS Human calicivirus Hu/NLV/GII/MD101-2/1987/US.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 OX NCBI_TaxID=159311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Hu/NLV/GII/MD101-2/1987/US;
 RX MEDLINE=21666333; PubMed=11807686;
 RA Green K.Y., Belliot G., Taylor J.L., Valdesuso J., Lew J.F., Kapikian A.Z., Lin F.Y.C.;
 RT "A Predominant Role for Norwalk-like Viruses as Agents of Epidemic Gastroenteritis in Maryland Nursing Homes for the Elderly."; J. Infect. Dis. 185:133-146 (2002).
 RL J. Infect. Dis. 185:133-146 (2002).
 DR EMBL: AY030312; RAK54360.1; --
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 SQ SEQUENCE 548 AA; 59898 MW; 776F74B4E58B6CC7 CRC64;

Query Match 97.2%; Score 2815; DB 12; Length 548;
 Best Local Similarity 96.2%; Pred. No. 1.4e-214;
 Matches 527; Conservative 15; Mismatches 6; Indels 0; Gaps 0;

QY 1 MKMSNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQOIIIDPWIMNPFVQAP 60
 DB 1 MKMSNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQOIIIDPWIMNPFVQAP 60
 QY 61 GGEFTVSPNSPGEVLLNLELGPENIPYLAHLARMYNGYAGGFVQVVLGNAFTAGKII 120
 DB 61 GGEFTVSPNSPGEVLLNLELGPENIPYLAHLARMYNGYAGGFVQVVLGNAFTAGKVI 120
 QY 121 FFAAIPNPFIDNLSAAQITMCCHVIVDVROLEPVLMPDVRNFFHYNQGSDSLRLIA 180
 DB 121 FFAAIPNPFIDNLSAAQITMCCHVIVDVROLEPVLMPDVRNFFHYNQGSDSLRLIA 180
 QY 181 MLYTPLRANSGDDVFTVSCRVLTPSPDFSNFLVPPVETSKTPFTLPILITISEMNS 240
 DB 181 MLYTPLRANSGDDVFTVSCRVLTPSPDFSNFLVPPVETSKTPFTLPILITISEMNS 240
 QY 241 RFPVPVIESLHTSPENIVVQCONGRVTLDGELMGTTQLLPISOICAFRGVLTSTGRASDQ 300
 DB 241 RFPVPVIESLHTSPENIVVQCONGRVTLDGELMGTTQLLPISOICAFRGVLTSTGRASDQ 300
 QY 301 ADTATPRLENYVHVLQDNLNGTTPDPAEDIPGLTDPFRGKVFVASQRNLDSTTRAH 360
 DB 301 ADTATPRLENYVHVLQDNLNGTTPDPAEDIPGLTDPFRGKVFVASQRNLDSTTRAH 360
 QY 361 EAKVDTTAGRTFKLGSLEISDSDDFQONQTKTTPVIGVYDNEAEFOQWSLPDYSQGF 420
 DB 361 EAKVDTTAGRTFKLGSLEISDSDDFQONQTKTTPVIGVYDNEAEFOQWSLPDYSQGF 420

QY 421 THNNMLAPAVANFPGEQLLFRSOLPSSGGRSNGVLDCLVPOEWWOHPYQESAPAQTOV 480
 DB 421 THNNMLAPAVANFPGEQLLFRSOLPSSGGRSNGVLDCLVPOEWWOHPYQESAPAQTOV 480
 QY 481 ALVRVYNPDGTGVLFEAKLHKLGFMTIANNGDSPITVPNGYFRFESWNPVYTLAPMGT 540
 DB 481 ALVRVYNPDGTGVLFEAKLHKLGFMTIANNGDSPITVPNGYFRFESWNPVYTLAPMGT 540
 QY 541 GNGRRRIQ 548
 DB 541 GNGRRRIQ 548

RESULT 9
 Q918A1 PRELIMINARY; PRT; 548 AA.
 AC Q918A1;
 DT 01-DEC-2001 (TREMELrel. 19, Created)
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Capsid protein.
 GN ORF2.
 OS Norwalk-like virus NLV/New Orleans/279/1994/US.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 OX NCBI_TaxID=171844;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Hu/NLV/New Orleans/279/1994/US;
 RX MEDLINE=97193806; PubMed=9041391;
 RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
 RT "A one-tube method of reverse transcription-PCR to efficiently amplify a 3-kilobase region from the RNA polymerase gene to the poly(A) tail of small round-structured viruses (Norwalk-like viruses)."; J. Clin. Microbiol. 35:570-577 (1997).
 RN [2]
 RP SEQUENCE OF 100-192 FROM N.A.
 RC STRAIN=Hu/NLV/New Orleans/279/1994/US;
 RX MEDLINE=98071277; PubMed=9407386;
 RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K., Seto Y., Monroe S.S., Glass R.I.;
 RT "Correlation of patient immune responses with genetically characterized small round-structured viruses involved in outbreaks of nonbacterial acute gastroenteritis in the United States, 1990 to 1995."; J. Med. Virol. 53:372-383 (1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Hu/NLV/New Orleans/279/1994/US;
 RX MEDLINE=20266071; PubMed=10804147;
 RA Ando T., Noel J.S., Fankhauser R.L.;
 RT "Genetic classification of 'Norwalk-like viruses.'"; J. Infect. Dis. 181:S336-S348 (2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Hu/NLV/New Orleans/279/1994/US;
 RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
 RT Submitted (AUG-2001) to the EMBL/GenBank/DBSJ databases.
 DR EMBL: AF414412; AAL2989.1; --
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 SQ SEQUENCE 548 AA; 59990 MW; F998C603777P8595 CRC64;

Query Match 97.1%; Score 2812; DB 12; Length 548;
 Best Local Similarity 96.4%; Pred. No. 2.4e-214;
 Matches 528; Conservative 12; Mismatches 8; Indels 0; Gaps 0;

QY 1 MKMSNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQOIIIDPWIMNPFVQAP 60
 DB 1 MKMSNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQOIIIDPWIMNPFVQAP 60
 QY 61 GGEFTVSPNSPGEVLLNLELGPENIPYLAHLARMYNGYAGGFVQVVLGNAFTAGKII 120

Db	61	GGFTVSPRNSPCEVLLNLELGEINPYLAHLARMYNGYAGGFVEQVVLGNAFTAGKYI	120
Qy	121	FAAIPNPFIDNLSAAQITMCPHVIVDQLEPVLNMPDVRNPFHYNQSDSLRLIA	180
Db	121	FAAIPNPFIDNLSAAQITMCPHVIVDQLEPVLNMPDVRNPFHYNQSDSLRLIA	180
Qy	181	MLYTPLRANNSGDDVFTVSCRVLTRPSDFSNFLVPPTVESKTKPFTLPILTISEMSNS	240
Db	181	MLYTPLRANNSGDDVFTVSCRVLTRPSDFSNFLVPPTVESKTKPFTLPILTISEMSNS	240
Qy	241	RPVPVIESLHTSPENIVVOCQNGRVTLTGELMGTTQLLPNOICAPRGTLTRSTNRASDQ	300
Db	241	RPVPVIESLHTSPENIVVOCQNGRVTLTGELMGTTQLLPNOICAPRGTLTRSTNRASDQ	300
Qy	301	ADTATPRLFNYYVHVDLNLNGTPYDPAEDI PCPLGTDPFRGKVGVSQRNLDSTTRAH	360
Db	301	ADTATPRLFNYYVHVDLNLNGTPYDPAEDI PCPLGTDPFRGKVGVSQRNLDSTTRAH	360
Qy	361	EAKVDTTGRFTPKGLSLEISTDDDFDQNPQTKFTFVGVGVDNEADFOQWSLPDYSGOF	420
Db	361	EAKVDTTGRFTPKGLSLEISTDDDFDQNPQTKFTFVGVGVDNEADFOQWSLPDYSGOF	420
Qy	421	THNNLAPAVAPNPPGQQLFFRSQLPSSGSRNGVLDCLVPOEWQHFYQESAPACTOV	480
Db	421	THNNLAPAVAPNPPGQQLFFRSQLPSSGSRNGVLDCLVPOEWQHFYQESAPACTOV	480
Qy	481	ALVRYVNPDTGKVLFEAKLHKLGFMTIANGSDSPITVPNGYFRFESWVNPFTYLPMTGT	540
Db	481	ALVRYVNPDTGKVLFEAKLHKLGFMTIANGSDSPITVPNGYFRFESWVNPFTYLPMTGT	540
Qy	541	NGRRRIQ 548	
Db	541	NGRRRIQ 548	
RESULT 11			
Qy	Q8V768	PRELIMINARY; PRT; 548 AA.	
AC	Q8V768	PRELIMINARY; PRT; 548 AA.	
DT	01-MAR-2002	(Tremblrel. 20, Created)	
DT	01-MAR-2002	(Tremblrel. 20, Last sequence update)	
DT	01-OCT-2003	(Tremblrel. 25, Last annotation update)	
DE	ORF2 protein (Capsid)		
OS	Norwalk virus, and		
OS	Norwalk-like virus.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;		
OC	Norovirus.		
OX	NCBI_TaxID=159312;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Hu/NLV/GII/MD134-10/1987/US;		
RK	MEDLINE=21666333; PubMed=11807686;		
RA	Green K.Y., Belliot G., Taylor J.L., Valdesuso J., Lew J.F.,		
RA	Kapikian A.Z., Lin F.Y.C.;		
RT	"A predominant role for Norwalk-like viruses as agents of epidemic		
RT	Gastroenteritis in Maryland nursing homes for the elderly."		
RL	J. Infect. Dis. 185:133-146 (2002).		
DR	EMBL; AY030313; AAK54361.1; -		
DR	InterPro; IPR004005; Calici.coat.		
DR	InterPro; IPR008975; Viral_cap.coat.		
DR	Pfam; PF00915; Calici.coat; 1.		
DR	SEQUENCE 548 AA; 59899 MW; OC02C5D953500351 CRC64;		
Query Match			
Best Local Similarity 97.0%; Score 2810; DB 12; Length 548;			
Matches 526; Conservative 16; Mismatches 6; Indels 0; Gaps 0;			
Qy	1	MKMASNDAPSDGAGLVPEINNEAMALDPVAGAAAPLTGQNIIDPWIMNPFVQAP	60
Db	1	MKMASNDAPSDGAGLVPEINNEAMALDPVAGAAAPLTGQNIIDPWIMNPFVQAP	60
Qy	61	GGFTVSPRNSPCEVLLNLELGEINPYLAHLARMYNGYAGGFVEQVVLGNAFTAGKYI	120

```
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 59949 MW; 589FC65F14B80F64 CRC64;

Query Match 96.8%; Score 2802; DB 12; Length 548;
Best Local Similarity 96.4%; Pred. No. 1.5e-213;
Matches 528; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQONIIDPWIMNNFVQAP 60
DB 1 MKMASNDAAPSNDGAAGLVPEINNEAMALEPVGAAIAAPLTGQONIIDPWIMNNFVQAP 60
QY 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGGFEVQVVLGNAGNAGTAKGII 120
DB 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGGFEVQVVLGNAGNAGTAKGII 120
QY 121 FAAPPPNPIDNLSAAQITMCPHVIVDVRLPEVNLPMPPDVNRNFFHYNQGSDSLRLIA 180
DB 121 FAAPPPNPIDNLSAAQITMCPHVIVDVRLPEVNLPMPPDVNRNFFHYNQGSDSLRLIA 180
QY 181 MLYTPLRANSGDDVFTVSCRVLTRPSDPFSFNFLVPPPTVESKTKPFTLPILTISEMNS 240
DB 181 MLYTPLRANSGDDVFTVSCRVLTRPSDPFSFNFLVPPPTVESKTKPFTLPILTISEMNS 240
QY 241 RPPVIESLHSTPTENIVVQCONGRVTLTGELMGTTQLLPISOICAFRGVLTTRSTGRASDQ 300
DB 241 RPPVIESLHSTPTENIVVQCONGRVTLTGELMGTTQLLPISOICAFRGVLTTRSTGRASDQ 300
QY 241 RPPVIDSLHSTPTENIVVQCONGRVTLTGELMGTTQLLPISOICAFRGVLTTRSTGRASDQ 300
DB 241 RPPVIDSLHSTPTENIVVQCONGRVTLTGELMGTTQLLPISOICAFRGVLTTRSTGRASDQ 300
QY 301 ADTATPRLFNYYVHVLQNLNGTVPYDPAEDIPGLGTPDFRGKVGVSQRNLDSTTRAH 360
DB 301 ADTATPRLFNYYVHVLQNLNGTVPYDPAEDIPGLGTPDFRGKVGVSQRNLDSTTRAH 360
QY 361 EAKVDTTACGRFTPKLGSLEISTDSDDFDQONQKFTPTVIGVGVNDAEFQOQMSLPYSGOF 420
DB 361 EAKVDTTACGRFTPKLGSLEISTDSDDFDQONQKFTPTVIGVGVNDAEFQOQMSLPYSGOF 420
QY 361 EAKVDTTSGRFTPKLGSLEISTDSDDFDQONQKFTPTVIGVGVNDAEFQOQMSLPYSGOF 420
DB 361 EAKVDTTSGRFTPKLGSLEISTDSDDFDQONQKFTPTVIGVGVNDAEFQOQMSLPYSGOF 420
QY 421 THNMMLPAVAPNPFGEQLLFFRSQLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQTV 480
DB 421 THNMMLPAVAPNPFGEQLLFFRSQLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQTV 480
QY 481 ALRVYVNDTGKVLPEAKLHKLGMFTIANNGDSPITVPNGYFRFESVWNPFTYLLPMGT 540
DB 481 ALRVYVNDTGKVLPEAKLHKLGMFTIANNGDSPITVPNGYFRFESVWNPFTYLLPMGT 540
QY 541 GNGRRRIQ 548
DB 541 GNGRRRIQ 548

RESULT 12
Q88291 PRELIMINARY; PRT; 548 AA.
AC Q88291;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Small round structured virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=37141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Auckland;
RA Regli W.J., Green D.H., Lewis G.D.;
RT "Outbreaks of oyster associated gastroenteritis: Virological
RT investigation and detection of a human Calicivirus.";
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U46039; AAB00437.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 59907 MW; 364BFA6EC263B6B2 CRC64;

Query Match 96.7%; Score 2799; DB 12; Length 548;
Best Local Similarity 95.6%; Pred. No. 2.6e-213;
Matches 524; Conservative 15; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQONIIDPWIMNNFVQAP 60
DB 1 MKMASNDAAPSNDGAAGLVPEINNEAMALEPVGAAIAAPLTGQONIIDPWIMNNFVQAP 60
QY 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGGFEVQVVLGNAGNAGTAKGII 120
DB 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGGFEVQVVLGNAGNAGTAKGII 120
QY 121 FAAPPPNPIDNLSAAQITMCPHVIVDVRLPEVNLPMPPDVNRNFFHYNQGSDSLRLIA 180
DB 121 FAAPPPNPIDNLSAAQITMCPHVIVDVRLPEVNLPMPPDVNRNFFHYNQGSDSLRLIA 180
QY 181 MLYTPLRANSGDDVFTVSCRVLTRPSDPFSFNFLVPPPTVESKTKPFTLPILTISEMNS 240
DB 181 MLYTPLRANSGDDVFTVSCRVLTRPSDPFSFNFLVPPPTVESKTKPFTLPILTISEMNS 240
QY 241 RPPVIESLHSTPTENIVVQCONGRVTLTGELMGTTQLLPISOICAFRGVLTTRSTGRASDQ 300
DB 241 RPPVIESLHSTPTENIVVQCONGRVTLTGELMGTTQLLPISOICAFRGVLTTRSTGRASDQ 300
QY 241 RPPVIDSLHSTPTENIVVQCONGRVTLTGELMGTTQLLPISOICAFRGVLTTRSTGRASDQ 300
DB 241 RPPVIDSLHSTPTENIVVQCONGRVTLTGELMGTTQLLPISOICAFRGVLTTRSTGRASDQ 300
QY 301 ADTATPRLFNYYVHVLQNLNGTVPYDPAEDIPGLGTPDFRGKVGVSQRNLDSTTRAH 360
DB 301 ADTATPRLFNYYVHVLQNLNGTVPYDPAEDIPGLGTPDFRGKVGVSQRNLDSTTRAH 360
QY 361 EAKVDTTACGRFTPKLGSLEISTDSDDFDQONQKFTPTVIGVGVNDAEFQOQMSLPYSGOF 420
DB 361 EAKVDTTACGRFTPKLGSLEISTDSDDFDQONQKFTPTVIGVGVNDAEFQOQMSLPYSGOF 420
QY 361 EAKVDTTSGRFTPKLGSLEISTDSDDFDQONQKFTPTVIGVGVNDAEFQOQMSLPYSGOF 420
DB 361 EAKVDTTSGRFTPKLGSLEISTDSDDFDQONQKFTPTVIGVGVNDAEFQOQMSLPYSGOF 420
QY 421 THNMMLPAVAPNPFGEQLLFFRSQLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQTV 480
DB 421 THNMMLPAVAPNPFGEQLLFFRSQLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQTV 480
QY 481 ALRVYVNDTGKVLPEAKLHKLGMFTIANNGDSPITVPNGYFRFESVWNPFTYLLPMGT 540
DB 481 ALRVYVNDTGKVLPEAKLHKLGMFTIANNGDSPITVPNGYFRFESVWNPFTYLLPMGT 540
QY 541 GNGRRRIQ 548
DB 541 GNGRRRIQ 548

RESULT 13
Q83880 PRELIMINARY; PRT; 548 AA.
ID Q83880
AC Q83880;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Capsid protein (Fragment).
OS Norwalk virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=11983;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SRV-OTH-25/89/J;
RX MEDLINE=94335115; PubMed=8057474;
RA Wang J., Jiang X., Madore H.P., Gray J., Desselberger U., Ando T.,
RA Seto Y., Oishi I., Lew J.F., Green K.Y., et al.;
RT "Sequence diversity of small, round-structured viruses in the Norwalk
RT virus group.";
RL J. Virol. 68:5982-5990 (1994).
DR EMBL: L23830; AAA59233.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
FT NON_TER 548
SQ SEQUENCE 548 AA; 59904 MW; 054379C8AA7541D9 CRC64;
```

RA Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995." J. Med. Virol. 53:372-383 (1997).
RL [3]
RN SEQUENCE FROM N.A.
RP STRAIN=Hu/NLV/Towson/313/1994/US;
RC MEDLINE=2026071; PubMed=10804147;
RX Ando T., Noel J.S., Fankhauser R.L.;
RA "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348 (2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Towson/313/1994/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glaes R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414414; AAL12995.1; -;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 5993 MW; 114F3907B3A26D89 CRC64;

Query Match 96.2%; Score 2786; DB 12; Length 548;
Best Local Similarity 96.0%; Pred. No. 2.8e-212;
Matches 526; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 1 MKMASNDAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQONIIDPWIMNFVQAP 60
Db 1 MKMASNDAPSNDGAAGLVPEINNEAMALEPVGAAIAAPLTGQONIIDPWIMNFVQAP 60
QY 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGVAGFEVQVVLGNAFTAGKII 120
Db 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGVAGFEVQVVLGNAFTAGKVI 120
QY 121 THNMNLAPAVAPNFPGEQLLFFRSQIPSSGGRNGVLDCLVPOEWVQHFYQSSAPAQTOV 480
Db 121 THNMNLAPAVAPNFPGEQLLFFRSQIPSSGGRNGVLDCLVPOEWVQHFYQSSAPAQTOV 480
QY 481 ALVRYVNPDTGKVLFEAKLHKLGFMTIANGDSPIITVPNGYFRFESWNPFTYLLAPMGT 540
Db 481 ALVRYVNPDTGKVLFEAKLHKLGFMTIANGDSPIITVPNGYFRFESWNPFTYLLAPMGT 540
QY 541 GNGRRRIQ 548
Db 541 GNGRRRIQ 548

RESULT 15
Q91722 PRELIMINARY; PRT; 548 AA.
ID Q91722
AC Q91722;

Query Match 96.5%; Score 2794; DB 12; Length 548;
Best Local Similarity 95.4%; Pred. No. 6.4e-213;
Matches 523; Conservative 16; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKMASNDAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQONIIDPWIMNFVQAP 60
Db 1 MKMASNDAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQONIIDPWIMNFVQAP 60
QY 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGVAGFEVQVVLGNAFTAGKII 120
Db 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGVAGFEVQVVLGNAFTAAKVI 120
QY 121 THNMNLAPAVAPNFPGEQLLFFRSQIPSSGGRNGVLDCLVPOEWVQHFYQSSAPAQTOV 480
Db 121 THNMNLAPAVAPNFPGEQLLFFRSQIPSSGGRNGVLDCLVPOEWVQHFYQSSAPAQTOV 480
QY 481 ALVRYVNPDTGKVLFEAKLHKLGFMTIANGDSPIITVPNGYFRFESWNPFTYLLAPMGT 540
Db 481 ALVRYVNPDTGKVLFEAKLHKLGFMTIANGDSPIITVPNGYFRFESWNPFTYLLAPMGT 540
QY 541 GNGRRRIQ 548
Db 541 GNGRRRIQ 548

RESULT 14
Q91725 PRELIMINARY; PRT; 548 AA.
ID Q91725
AC Q91725
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Towson/313/1994/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171846;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Towson/313/1994/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577 (1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Towson/313/1994/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,

DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Brattleboro/321/1995/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Brattleboro/321/1995/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Brattleboro/321/1995/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
RA Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995.";
RL J. Med. Virol. 53:372-383(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Brattleboro/321/1995/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk'-like viruses.";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Brattleboro/321/1995/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.L., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414415; AAL12998.1; -;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 55933 MW; 7B8AFF9AF1469158 CRC64;

Query Match 96.0%; Score 2781; DB 12; Length 548;
Best Local Similarity 95.8%; Pred. No. 6.9e-212;
Matches 525; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAPLTGOQNIIDPWIMNNFVOAP 60
DB 1 MKMASNDAAPSNDGAAGLVPEINNEAMALEPVAGAAIAPLTGOQNIIDPWIMNNFVOAP 60

QY 61 GGEFTVSPRNSPGEVLLNLELGEINPYLAHLARMYNGYAGGFEVQVVLGNAFTAGKII 120
DB 61 GGEFTVSPRNSPGEVLLNLELGEINPYLAHLARMYNGYAGGFEVQVVLGNAFTAGKVI 120

QY 121 FAAPPPNPIDNLSAAQITMCPHVIVDVRQLEPVNLPMPPDVNRNFFHYNQGSDSLRLIA 180
DB 121 FAAPPPNPIDNLSAAQITMCPHVIVDVRQLEPVNLPMPPDVNRNFFHYNQGSDSLRLIA 180

QY 181 MLYTFLRANSGDDVFTVSCRVLTPSPDFSENFVLPPTVESKTKPFTLPILITISEMNS 240
DB 181 MLYTFLRANSGDDVFTVSCRVLTPSPDFSENFVLPPTVESKTKLFTLPILITISEMNS 240

QY 241 RPPVPIESLHSTPTENIVVQCONGRVTLTGELMGTTQLLPQICAFRGVLTSTSRASDQ 300
DB 241 RPPVPIESLHSTPTENIVVQCONGRVTLTGELMGTTQLLPQICAFRGVLTSTSRASDQ 300

QY 301 ADTATPRLFNYYVHVLQNLNGTTPYDPAEDIPGLGTPDFRGKVFQVASQRNLDSTTRAH 360
DB 301 ADTATPRLFNYYVHVLQNLNGTTPYDPAEDIPGLGTPDFRGKVFQVASQRNLDSTTRAH 360

DB 301 ADTPTPRLFNHRWHIQLDNLNGTTPYDPAEDIPAPLGTPDFRGKVFQVASQRNPDSTTRAH 360
QY 361 EAKVDTTAGRTPTPKLGSLEISTDSDDFQONQPTKFTPVGIGVDNEAEFQWLSLPDYSGQF 420
DB 361 EAKVDTTSDRPTPKLGSLEIITESGDFDTNQSTKFTTPVGIGVDNEAEFQWLSLPNYSGQF 420
QY 421 THNMNLAPAVAPNFFPGEQLLFFRSQLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQTOV 480
DB 421 THNMNLAPAVAPNFFPGEQLLFFRSQLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQTOV 480
QY 481 ALVRYVNPDTGKVLPEAKLHKLGMWTIANNGDSPITVPNGYFRFESWNPFTYLPAMGT 540
DB 481 ALVRYVNPDTGKVLPEAKLHKLGMWTIANNGDSPITVPNGYFRFESWNPFTYLPAMGT 540
QY 541 GNGRRRIQ 548
DB 541 GNGRRRIQ 548

Search completed: June 1, 2004, 13:53:17
Job time : 32.9054 secs

THIS PAGE BLANK (USPTO)

Result No.	Score	Query			DB	ID	Description
		Match	Length	§			
1	2034	70.2	542	2	S60616	capsid protein - h	
2	1947.5	67.2	539	2	S40111	capsid protein - h	
3	1211	41.8	530	2	B37471	capsid protein - N	
4	1123.5	38.8	546	2	B37491	major capsid prote	
5	302.5	10.4	576	2	A53982	capsid protein - E	
6	284.5	9.8	2344	1	RRWWRH	genome polyprotein	
7	281.5	9.7	702	1	A48562	coat protein - San	
8	277.5	9.6	2344	2	S55399	genome polyprotein	
9	276	9.5	2344	2	S64740	genome polyprotein	
10	275.5	9.5	671	1	VCWWF9	coat protein - fel	
11	269	9.3	703	1	C48562	coat protein - San	
12	262	9.0	668	1	VCWVFF	coat protein - fel	
13	261.5	9.0	668	1	VCWVFC	coat protein - fel	
14	257	8.9	668	2	JQ2354	capsid protein - f	
15	255	8.8	668	2	JQ2356	capsid protein - f	
16	138.5	4.8	4936	2	AH2515	hypothetical prote	
17	137.5	4.7	2194	1	GNVYE7	genome polyprotein	
18	133.5	4.6	3085	2	T00327	polyprotein - infe	
19	128.5	4.4	2205	1	GNVY2W	genome polyprotein	
20	126.5	4.4	2207	1	GNVY5P	genome polyprotein	
21	126	4.4	1965	2	S75200	fat protein - Syne	
22	125	4.3	2175	1	GNVYBE	genome polyprotein	
23	124.5	4.3	940	2	D89723	protein F39D8.1b [
24	124.5	4.3	945	2	T21998	hypothetical prote	
25	121.5	4.2	929	2	A44048	genome polyprotein	
26	121	4.2	2233	2	B95075	beta-galactosidase	
27	120.5	4.2	943	2	JC4081	sucrase/fructanase	
28	119	4.1	2332	1	GNVYF	genome polyprotein	
29	118.5	4.1	1023	2	T48997	epsin-like protein	

Db 349 EPANRGHDAVPTTYTAQYTPKLGQIQIGTWQDDLTVNQPVKFTPVGL--NDTEHFNQW 405

QY 412 SLDPYSGQFTHNMNLAPAVAFNPGEQLLFRSQLPSSGGRSGVLDCLVPQEWQHFYQ 471
Db 406 VVPRVAGALNLTNLAPSVAVPFGERLLFRSHLPLKGGYGNPAIDCLLPQEWQHFYQ 465
QY 472 ESAPQATQVALVRYNPDGTGKVLPAKHLKGLFMTIANNNGDSPIVTPNGYFRFRESWNP 531
Db 466 EAAPSMSEVALVRYNPDGTGKVLPAKHLKGLFMTIANNNGDSPIVTPNGYFRFRESWNP 525
QY 532 FYTLAPMGTCNGRRRIQ 548
Db 526 FYSLAPMGTCNGRRRVQ 542
RESULT 2
S40111
capsid protein - human calicivirus (strain Bristol isolate B493)
C:Species: human calicivirus
A:Variety: strain Bristol isolate B493
C:Date: 25-Dec-1994 #sequence_revision 27-Feb-1997 #text_change 28-Jul-2000
C:Accession: S40111
R:Green, S.M.; Dingle, K.E.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
submitted to the EMBL Data Library, December 1993
A:Description: Human enteric caliciviridae: a new prevalent SRSV group defined by RNA-de
A:Reference number: S40111
A:Accession: S40111
A:Molecule type: genomic RNA
A:Residues: 1-539 <GRE>
A:Cross-references: EMBL:X76716; NID:G436410; PIDN:CAAS4134.1; PID:G436411
A:Experimental source: human enteric calicivirus strain Bristol isolate B493
A:Superfamily: human calicivirus capsid protein
C:Keywords: capsid protein; coat protein
Query Match 67.2%; Score 1947.5; DB 2; Length 539;
Best Local Similarity 66.1%; Pred. No. 1.1e-138;
Matches 364; Conservative 74; Mismatches 94; Indels 19; Gaps 4;
QY 1 MKMASNDAAPSNDGAAG---LVPEIN-NEAMALDPVAGAAIAAPLTGQOQNIIDPWIMNNFVQAP 60
Db 1 MKMASNDANPSDGAANLVPEINNEAMALDPVAGAAIAAPLTGQOQNIIDPWIMNNFVQAP 60
QY 61 GGEFTVSPRNSPGEVLLNLELGPENIPYLAHLARMYNGYAGGFEVQVVLGNAFTAGKII 120
Db 61 GGEFTVSPRNPAGEILWSALPGDLNLYLHLSRMVNGYAGGFEVQVVLGNAFTAGKVI 120
QY 121 FAALPPNFPIDNLSAAQITMCPHVIDVROLEPNLMPDVRNFFHYNQGSRLRLIA 180
Db 121 FAAVPPNFPTEGLSPSQVTWPHIIVDVROLEPNLMPDVRNFFHYNQANDSTLKLIA 180
QY 181 MLYTPLRANNSGDDVFTVSCRVLTRPSDFSNFLVPTVESKTKPFTLPILATISEMSNS 240
Db 181 MLYTPLRANNGDDVFTVSCRVLTRPSDFSNFLVPTVESKTKPFTVPVLTVEEMSNS 240
QY 241 RFPVPIESLHTSPTEINIVVQCQGRVTLDELMTGTLQLPSQICAFRGVLTSTRASDQ 300
Db 241 RFPVPLEKLYGPSSAFVVPQNGKCTDGLVLLGTQLSAINICNFRGDVTHIAG----- 295
QY 301 ADTATPRLFNYYHVLQDNLNGTPLYDAEDIPGLTDPDRGKVGVASQ-RNLDSTTRA 359
Db 296 -----SHDYTMNLASQNSNYDPTTEIPAPLGTDPDFVGKIQGLLTQTTRADGSTRA 346
QY 360 HEAKVDTTTRAEKAVDTTTRAGRTPKLGSLEISTDSDDFQDQPTKFTVPGIGVDN----- 404
Db 347 HKATVSTGSVHTPKLGSVQVTTDNDNFQAGQNTKFTVPGVIGQGDHQQNEPQWLLFN 406
QY 416 YSQQFTHNMNLAPAVAFNPGEQLLFRSQLPSSGGRSGVLDCLVPQEWQHFYQ 475
Db 407 YSGRTGHNVLAPAVAPTGPGEQLLFRSTWPGCGSGYPNMNLDCLLPQEWLHLYQGAAP 466
QY 476 ACTQVALVRYNPDGTGKVLPAKHLKGLFMTIANNNGDSPIVTPNGYFRFRESWNPFTYL 535
Db 467 AOSDVALKRVNPDGTGKVLPAKHLKGLFMTIANNNGDSPIVTPNGYFRFRESWNPFTYL 526
QY 536 APMGTGNGRRR 546

Db 527 APMGTGNGRRR 537
|||||
RESULT 3
B37471
capsid protein - Norwalk virus
C:Species: Norwalk virus
C:Date: 24-Feb-1994 #sequence_revision 25-Apr-1997 #text_change 28-Jul-2000
C:Accession: B37471
R:Jiang, X.; Wang, M.; Wang, K.; Estes, M.K.
Virology 195, 51-61, 1993
A:Title: Sequence and genomic organization of Norwalk virus.
A:Reference number: A37471; MUID:93303939; PMID:8391187
A:Accession: B37471
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: genomic RNA
A:Residues: 1-530 <JIA>
A:Cross-references: GB:M87661; NID:g1061311; PIDN:AAB50466.1; PID:g1061313
A:Note: sequence extracted from NCBI backbone (NCBIP:134157)
C:Superfamily: human calicivirus capsid protein
Query Match 41.8%; Score 1211; DB 2; Length 530;
Best Local Similarity 45.4%; Pred. No. 2.9e-83;
Matches 260; Conservative 74; Mismatches 161; Indels 78; Gaps 11;
QY 1 MKMASNDAAPSNDGAAG---LVPEIN-NEAMALDPVAGAAIAAPLTGQOQNIIDPWIMNNF 56
Db 1 MMWASKATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIMNNF 60
QY 57 VOAPGGEFTVSPRNSPGEVLLNLELGPENIPYLAHLARMYNGYAGGFEVQVVLGNAFTA 116
Db 61 VOAPGGEFTTSPNTPGDVLFDSLGLHLPFLHLSQMYNGVWGNRVRIMLAGNAFTA 120
QY 117 GKIIIFAAIPNFPIDNLSAAQITMCPHVIDVROLEPNLMPDVRNFFHYNQGSRL 176
Db 121 GKIIIVSCIPPPGFGSHNLIIAQTLLFPFIADVTLDIEVPLEDVRNVLPHNDRNQOTM 180
QY 177 RLIAMLYTLRANNSGDDVFTVSCRVLTRPSDFSNFLVPTVESKTKPFTLPILATISE 236
Db 181 RLVCMLYTLRATGGTGDGDFVAGRVMTCTSPDNFLFLVPTVEQKTRPFTLPNPLSS 240
QY 237 MSNRFVPIESLHTSPTEINIVVQCQGRVTLDELMTGTLQLPSQICAFRGVLTSTR 296
Db 241 LSNRAPLPISSIGISPDNVQSVQFQNGKCTDDELVGTTPVSLSHVAKIRG-----TSN 295
QY 297 ASDQADTATPRLFNYYHVLQDNLNGTPLYDAEDIPGLTDPDRG-----KVFGVAS 349
Db 296 GT-----VINLTDLDTGTFHPFEG-PADIGPFDLGGCDWHNMNTQFGHSS 339
QY 350 QRNLSTTRAEKAVDTTTRAGRTPKLGSLEISTDSDDFQDQPTKFTVPGIGVDN----- 404
Db 340 QTQYD-----VDTTPTDFVPHLGSIQAN-----GIGSGNYGVGL 373
QY 405 -----EAFQOWSLPDYSGQFTHNMNLAPAVAFNPGEQLLFRSQLPSSGGRSN 454
Db 374 SWISPPSHPSGQVDLWKIPNYGSSITATHLAPSVYPPGFEVLVFFMSXMGPGCAYN- 432
QY 455 GVLDCLVPQEWQHFYQESAPATQVVALVRYNPDGTGKVLPAKHLKGLFMTIANNNGDS- 513
Db 433 --LPCLLPQEIYSHLASEQAQPTVGEAALLHYVDPDTRNLGFEKAYPDGFLTCVPGNGASS 490
QY 514 -PIVTPNGYFRFRESWNPFTYLAPMGTCNGRR 545
Db 491 GPQQLPINGVVFVSVWSRFPYQLKPVGTASSAR 523
RESULT 4
B37491
major capsid protein [similarity] - Southampton virus
N:Alternate names: orf2 protein
C:Species: Southampton virus
C:Date: 22-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 18-Aug-2000

338 PDPRGKGVGVASQNRDLSTTRAHEAKVDVT-----AGRFTPKGLSLEISTDSDDFDQN 390
454 POTTITAMTASNGVDYVAEYRITNNGTHPKGFVIMGNLTTKV-----KGSNDLGET 507
391 QPTK---FTPVG 399
508 QQTSTRTLFASVG 519
RESULT 8
S55399
genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)
C;Species: rabbit hemorrhagic disease virus
A;Variety: isolate BS89
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 18-Jun-1999
C;Accession: S55399
R;Roessl, C.
submitted to the EMBL Data Library, May 1995
A;Reference number: S55399
A;Accession: S55399
A;Molecule type: genomic RNA
A;Residues: 1-2344 <ROS>
A;Cross-references: EMBL:X87607; NID:g854640; PIDN:CAA60910.1; PID:g854641
A;Experimental source: isolate BS89
C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C;Keywords: polyprotein
Query Match 9.6%; Score 277.5; DB 2; Length 2344;
Best Local Similarity 27.0%; Pred. No. 4.6e-12;
Matches 113; Conservative 57; Mismatches 174; Indels 75; Gaps 20;
QY 9 APSNDGA-----AGLVPEINNEAMALDP-----VAGAAATAPLTGOONI 47
DB 1773 APQCEAAGTATTASVPGTTTD--GLDFGVVATTSVVTAENSSASIAAGIGGP-PQQVDQ 1829
QY 48 IDPWIMNFVQAPCGEFTVSPNSPGEVLLNLELGPINPILAHARMYNGYAGGFVQV 107
DB 1830 QETWRTNFY---NDVFTWSVADAPGSILYVQHSPPQNNPFTAVLSQMYAGWAGMQFRF 1886
QY 108 VLAGNAFTAGKIIFAAIPPNPFDNLNSAAQITMCQPHVIVDVRQLEPNVLPMPDVRRNFFH 167
DB 1887 IVAGSGVGGRLVAIVPPGIEG--PGLEVRQPHVVIDARSLEPVTITMPDLRPNMYH 1944
QY 168 YNQGSDSRLRLIAMLPTPLRANSGDDVFTVSCVRLTRSPDRSFNVLPP---TVESKT 224
DB 1945 PTGDPGLVPTLVLSVYNL-INPFGGTSATQVTRPSDDFFVIRAPSSKTVDIS 2003
QY 225 KP--FTLPILTISEMSNSRFPVPIESLHTSPENIVVQCNQGRVTLDELMTGTTQLLP 282
DB 2004 PAGLITTPVLT-GVGNDRNWNGQIVGL--QVPVGGFSTC-NRHWNLNGSTYGMSSSPRAD 2059
QY 283 ICAPRGVLTSTRASDAQDTATPLRFNYVHVQDLNLTGTPYDPAEDIPGLGTPD--- 339
DB 2060 IDHRKG-----SASYPGNSATNVL--QFWYA---NAGSAVDNPISQV-APDGPDMSF 2106
QY 340 ---FRGK-----VGVASQNRDL-----STTRAHEAKVDVTAGRTFKGLSLEISTDS 384
DB 2107 VPFNGPGIPAAAGWGVGAIWNSGAPNVTVQAYE-----LGFAATGAPGNLQPTTNT 2159
RESULT 9
S64740
genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)
N;Contains: VP60 protein
C;Species: rabbit hemorrhagic disease virus
A;Variety: isolate AST/89
C;Date: 12-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Jun-2000
C;Accession: S64740; S46944; S49018; S65012
R;Casals, R.; Martin-Alonso, J.; Boga, J.; Parra, F.
submitted to the EMBL Data Library, May 1995
A;Description: Genomic organization of rabbit hemorrhagic disease virus determined by d
A;Reference number: S64740
A;Accession: S64740

1829 QETWRTNFY---NDVFTWSVADAPGSILYVQHSPPQNNPFTAVLSQMYAGWAGMQFR 1885
107 VVLAGNAFTAGKIIFAAIPPNPFDNLNSAAQITMCQPHVIVDVRQLEPNVLPMPDVRRNFF 166
1886 FIVAGSGVGGRLVRAVIPPGEIG--PGLEVRQPHVVIDARSLEPVTITMPDLRPNMY 1943
167 HYNQGSDSRLRLIAMLPTPLRANSGDDVFTVSCVRLTRSPDRSFNVLPP---TVESK 223
1944 HFTGDPGLVPTLVLSVYNL-INPFGGTSATQVTRPSDDFFVIRAPSSKTVDIS 2002
224 TKP--FTLPILTISEMSNSRFPVPIESLHTSPENIVVQCNQGRVTLDELMTGTTQLLP 281
2003 SPAGLLITTPVLT-GVGNDRNWNGQIVGL--QVPVGGFSTC-NRHWNLNGSTYGMSSSPRFG 2058
282 QICAFRGVLTSTRASDAQDTATPLRFNYVHVQDLNLTGTPYDPAEDIPGLGTPD--- 339
2059 DIDHRGGSASYSGSNATNVLQ-----FWYANA-----GSAIDNPISQVAPDGFDM 2105
340 ---FRGK-----VGVASQNRDL-----STTRAHEAKVDVTAGRTFKGLSLEISTDS 384
2106 FVPFNGPGIPAAAGWGVGAIWNSGAPNVTVQAYE-----LGFAATGAPGNLQPTTNT 2159
RESULT 7
A48562
coat protein - San Miguel sea lion virus (serotype 1)
N;Alternate names: capsid protein
C;Species: San Miguel sea lion virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C;Accession: A48562
R;Neill, J.D.
Virus Res. 24, 211-222, 1992
A;Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel s
eins.
A;Reference number: A48562; MUID:92410750; PMID:1529644
A;Accession: A48562
A;Molecule type: genomic RNA
A;Residues: 1-702 <NEI>
A;Cross-references: GB:M87481; NID:g334882; PIN:AAAL6217.1; PID:g334884
A;Note: sequence extracted from NCBI backbone (NCBIN:113564, NCBI:113565)
C;Superfamily: feline calicivirus coat protein
C;Keywords: capsid protein; coat protein; glycoprotein
F;208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 9.7%; Score 281.5; DB 1; Length 702;
Best Local Similarity 25.9%; Pred. No. 4e-13;
Matches 112; Conservative 57; Mismatches 160; Indels 103; Gaps 19;
QY 8 AAPNDGAAGLVPEINNEAMALDPV---AGAAATAPLTGOONIIDPWIMNFVQAPGGE 63
DB 151 AESDGPGGADIVTEQGTIVQQQPVPAQSALTTLAAASTGK--TVD-----CE 196
QY 64 FTV-----SPRNSPGEVLLNLELGPINPILAHARMYNGYAGGFVQVVLGNA 113
DB 197 WTTFFSYHTAVNWSSTEAQKILFSRALSPELNPYLRHISLSYTWSSGDIIVRTVSGSG 256
QY 114 FTAGKIIFAAIPPNPFDNLNSAAQITMCQPHVIVDVRQLEPNVLPMPDVRRNFFHYNQSD 173
DB 257 VFGGKLAALIVPPG--IEFVESPTMLQYPHVLFDAQRTPEVIFTIPDIRKTLHSMDDTD 314
QY 174 SRLRLIAMLPTPL-RANNSGDDVFTVSCVRLTRSPDRSFNVLPPVPTVESKTPFTLPIL 232
DB 315 T-TRUIVIMVYNELINPYEQSEPKSSCSITVETRSSDFTSLKPP--GSLLKHSIP-- 369
QY 233 TISEMSNSRFPVPIESLH-----TSPTENIVVQCNQGRVTLDELMTGTTQLLSQICAP 286
DB 370 --SDL-----IPRNSRHMGNRWNSTIDGFWQVR----- 397
QY 287 RGVLTSTRASDAQDTATPLRFNY--WHVOLDNLN-GTPYDPAED-----IPG-PLGT 337
DB 398 ----VFQSNRHFDFDSTTTGWSTPTPIPIEVLTKLDKGGQYFKYTDTEKSLVPLGLPDGW 453

A:Molecule type: genomic RNA
A:Residues: 1-2344 <CAG>
A:Cross-references: EMBL:249271; NID:g1182032; PIDN:CAA89265.1; PID:g1150552
A:Experimental source: isolate AST/89
R:Boğa, J.; Casais, R.; Martin, M.; Martín-Alonso, J.; Carmenes, R.; Prieto, M.; Parra, F.
submitted to the EMBL Data Library, July 1993
A:Description: Molecular cloning, sequence and expression of the capsid protein gene from
A:Reference number: S46944
A:Accession: S46944
A:Molecule type: genomic RNA
A:Residues: 1650-2344 <BOG>
A:Cross-references: EMBL:224757; NID:g515622; PIDN:CAA80881.1; PID:g515623
A:Experimental source: isolate AST/89
R:Parra, F.; Boğa, J.A.; Martin, M.S.; Casais, R.
Virus Res. 27, 219-228, 1993
A:Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus supp
A:Reference number: S49018; MUID:93255896; PMID:8488721
A:Accession: S49018
A:Molecule type: genomic RNA
A:Residues: 1650-1796 <PAR>
A:Cross-references: EMBL:224757
A:Experimental source: isolate AST/89
A:Accession: S65012
A:Molecule type: protein
A:Residues: 1767-1779; 1875-1877, 'X', 1879-1881, 1936-1938, 'X', 1940-1941 <PAM>
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: blocked amino end; polyprotein
Query Match 9.5%; Score 276; DB 2; Length 2344;
Best Local Similarity 22.8%; Pred. No. 6e-12;
Matches 109; Conservative 60; Mismatches 161; Indels 148; Gaps 19;
QY 9 APSNDGA---AGLVEINNEAMALDP-----VAGAAIAAFLTGQNI 47
DB 1773 APOGEAAGTATTASVGTITDGM--DPGVATTSVVTAESSASIAATAGIGGP-PQVDQ 1829
QY 48 IDPWNNVQAPGGEFTVSPRNSPGEVLNLNLELGPINPYLAHARMYNGVAGFEVQV 107
DB 1830 QETWRTNFY---NDVFTWSVADAPGSIYTVQHSPPNNPFTAVLSQMYAGWAGGMQFRF 1886
QY 108 VLAGNAFTAGKIFRAIPNFPIDNLSAAQITMCCHVIVDVQLEPNLPMDDVRNPF 167
DB 1887 IVAGIGVFGRLVAIVPGIEIG--PGLEVQFPFHVIDARSLPEVTITMPDLRPNMYH 1944
QY 168 YNQSDSRLRLIAMLYPLRANNSGDVFTVSCRVLTRPSDFSNFLVPP---TVESKT 224
DB 1945 PTGDCGLVTLVLVYNNL-INPFGGSTSAIQVTVETRSEDEFVWIRAPSKTVDSLS 2003
QY 225 KP--FTLPILTISEMSNSRFPVPIESLHTSPTENIVVOCONGRVTLDBELMGTQLLPSQ 282
DB 2004 PAGLLTTPVLT-GVGNDNRW-----NGQIVG-LQVPQG- 2034
QY 283 ICAFRGVLTSTRASDAQDTATPRLFNYVHVDNLNGTPYDPAEDIPGLTPTDFRG 342
DB 2035 -----GFST-----CNRHW-----NLNGSTY----- 2050
QY 343 KVFVGVAQRNLSDSTTRAHEAKVDTTAGRTPTPKLSLEISTSDSDPDQNPQTK-----FTP 397
DB 2051 -----GWSPPFGDIGHRGASYPGNATVQLQFWYAN 2084
QY 398 VGIGVDNE-AEFQWQSLPYSGQFTNHNMLAPAVAPNPFEGQLLFFRSQLPSSGGRSN 454
DB 2085 AGSAIDNPISQVAPDFFD-----MSFVPFNGPGIIPAAAGWVGFGAIWNSGAPN 2134
RESULT 10
VCWMF9
coat protein - feline calicivirus (strain F9)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 28-Jul-2000
C:Accession: B43382; C45538; PQ0407; S23702
R:Carter, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.; Turner, P.C.

Virology 190, 443-448, 1992
A:Title: The complete nucleotide sequence of a feline calicivirus.
A:Reference number: A43382; MUID:92410623; PMID:1529544
A:Accession: B43382
A:Molecule type: genomic RNA
A:Residues: 1-671 <CAR1>
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879
R:Carter, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.
Arch. Virol. 122, 223-235, 1992
A:Title: Identification and sequence determination of the capsid protein gene of feline
A:Reference number: A45538; MUID:92117861; PMID:1731695
A:Accession: C45538
A:Molecule type: genomic RNA, protein
A:Residues: 1-671 <CAR2>
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879
A:Experimental source: strain F9
A:Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)
R:Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.
J. Gen. Virol. 73, 2429-2433, 1992
A:Title: The cloning, sequencing and expression of a major antigenic region from the fel
A:Reference number: PQ0407; MUID:93019069; PMID:1402818
A:Accession: PQ0407
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 217-266, 'SI', 269-336, 'V', 338-395, 'S', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>
A:Cross-references: PIDN:AA823553.1; PID:g257083
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 9.5%; Score 275.5; DB 1; Length 671;
Best Local Similarity 22.3%; Pred. No. 1.1e-12;
Matches 123; Conservative 70; Mismatches 174; Indels 185; Gaps 23;
QY 67 SPRNSPGEVLLNLELGPINPYLAHARMYNGVAGFEVQVVLGAGNAFTAGKIIPAAIPP 126
DB 179 STSETQKILFKQSGLGPLLNPYLEHLAKLYVAMSGSIEVRPSISGSGVFGKLAIVVPP 238
QY 127 NPPIDNLSAAQITMCCHVIVDVQLEPNLPMDDVRNPFNYNQSDSRLRLIAMLYTLP 186
DB 239 G--VDPVQSTSMLOYPHVLFDARQVEPVIFCLPDLRLSTLYHLMSDITD--TSLVIMVYNDL 295
QY 187 RANNSGDDVFTVSC--RVLTSPSPFSNFLVPP-----TVES-----KT----- 224
DB 296 -INPYANDANSSGCIIVTVETKPGDPKFKHLLKPPGSMTHGSIPLDIPKTSLSMIGNRY 354
QY 225 -----KPFTL-----PI-LTISEMSNSRFPVPIESLHTSP 253
DB 355 WSDITDFVIRPEVFOANRHFDFNQETAGWSTFRPRPISVTITEQNGAKLIGV-----A 408
QY 254 TENIVVQCQNG--RVTLDGELMGTQLLPSQICAFRGVLTSTRASDAQDTATPRLFNY 311
DB 409 TDYIVPGIDPGWPDITIGEL-----IPAGDYAITNGTN----- 443
QY 312 YMHVQLDNLNGTPYDPAEDIPGLTPTDFRGKVGKVGASQRNLSDSTTRAHEAKVDTTAGRF 371
DB 444 -----DITTATGYDTADII--NNTNFRGMVYICGSLQRAMG-----DKKISNTAFIT 488
QY 372 TPKLSLEISTSDP-----PDQNPQTKFTPVGIGVDNEAEFQWQSLPYSGQFTH 422
DB 489 TATL-----DGDNNKINPCNTDQSKIVVFDNHNHGKKAQTSDDTLALLGYTG----- 537
QY 423 NNNLAPAVAPNPPGQLL-----FFRSQLPSSGGRSNGVLDCLVPQEWQOH--FYQSS 473
DB 538 -----IGEQAIGSDRDRVVRISTLPETGARG-----NHPIFYKNS 573
QY 474 AP-----AQTQVALRVYVNP-----DTGKVLFEAKLHLKGFMT 506
DB 574 IKLGVIIRSIDVFNSQLHTSRQLSLNHYLLPPDSFAVVIIDNSGWNFDIGSDGFSF 633
QY 507 IANNGDSITVP 518
DB 634 VGVSGFGKLEFP 645

C;Keywords: capsid protein; coat protein; glycoprotein
F;177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.0%; Score 262; DB 1; Length 668;
Best Local Similarity 34.4%; Pred. No. 1.1e-11;
Matches 54; Conservative 32; Mismatches 59; Indels 12; Gaps 4;

Qy 67 SPNSPGEVLLNLELGPENIPYLAHLARMYNGAGVGVVVLGNAGNAFTAGKIIFAAIPP 126
Db 179 STSETQKILFKQSLGFLNPNYTHLAKLYVWSSGVDFRFSISGSGVGGKLAALVWPP 238
Qy 127 NFPIIDLSAAQITMCPHVIVDVQLSPVNLPMDFVNNFFHYNQSGDSRLRLIAMLYTPL 186
Db 239 G--IDPQVSTMLQYPHVLFDAQVPEVIFSDLRSTLYHLSMDTDT-TSLVIMVYNDL 295
Qy 187 -----RANNSGDDVFTVSCRVLTSPDPSFNFVLP 218
Db 296 INPYANDNSGGCIIVTVE----TKPGDFKFKHLLKPP 328

RESULT 13
VCWVFC
coat protein - feline calicivirus (strain Japanese F4)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C:Accession: B40481
R:Tohya, Y.; Taniuchi, Y.; Takahashi, E.; Utagawa, E.; Takeda, N.; Miyamura, K.; Yamaza
Virology 183, 810-814, 1991
A:Title: Sequence analysis of the 3'-end of feline calicivirus genome.
A:Reference number: A40481; MUID:91306470; PMID:11853578
A:Accession: B40481
A:Molecule type: genomic RNA
A:Residues: 1-668 <TOH>
A:Cross-references: GB:D50357; NID:g221264; PIDN:BAA14371.1; PID:g221266
C:Superfamily: feline calicivirus coat protein
C;Keywords: capsid protein; coat protein; glycoprotein
F;177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.0%; Score 261.5; DB 1; Length 668;
Best Local Similarity 30.3%; Pred. No. 1.2e-11;
Matches 59; Conservative 41; Mismatches 78; Indels 17; Gaps 6;

Qy 30 DPVAGAAIAAPLTGQONTIDPW-INNFVQAPGGEFTVSPNSPGEVLLNLELGPENPY 88
Db 145 EPSAQWSTAADMASGKSVDSWEAFSPHTS----VNMSTSETQKILFKQSLGFLNPNY 200
Qy 89 LAHLARMYNGAGVGVVVLGNAGNAFTAGKIIFAAIPNFPIDNLSAAQITMCPHVIVDV 148
Db 201 LEHLSKLYVWSSGSTEVRFSISGSGVGGKLAALVWPPG--VDPQVSTMLQYPHVLFDA 258
Qy 149 ROLPEVNLPMDFVNNFFHYNQSGDSRLRLIAMLYTPL-----RANNSGDDVFTVSCRVL 203
Db 259 RQVEPVIFIPDLRSTLYHWSMDTDT-TSLVIMVYNDLINPYANDNSGGCIIVTVE---- 313
Qy 204 TRPSDPSFNFVLP 218
Db 314 TKPGDFKFKHLLKPP 328

RESULT 14
JCQ354
capsid protein - feline calicivirus (strain NADC)
C:Species: feline calicivirus
C>Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: JQ2354
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable
A:Reference number: JQ2354; MUID:94065683; PMID:7504075
A:Accession: JQ2354
A:Molecule type: mRNA
A:Residues: 1-668 <SEA>

RESULT 11
C48562
coat protein - San Miguel sea lion virus (serotype 4)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: C48562
R:Neill, J.D.
Virus Res. 24, 211-222, 1992
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea
eins.
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Accession: C48562
A:Molecule type: genomic RNA
A:Residues: 1-703 <NE1>
A:Cross-references: GB:M87482; NID:g334886; PIDN:AAAL6220.1; PID:g334888
A>Note: sequence extracted from NCBI backbone (NCBIP:1113567)
C:Superfamily: feline calicivirus coat protein
C;Keywords: capsid protein; coat protein; glycoprotein
F;89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.3%; Score 269; DB 1; Length 703;
Best Local Similarity 32.3%; Pred. No. 3.5e-12; Indels 18; Gaps 7;
Matches 71; Conservative 38; Mismatches 89;

Qy 8 AAPSNDGAAGLVE-----INNEAMALDPVAGAAIAAPLTGQONTIDPWNNFVQAPGGE 63
Db 151 AESDGFSAIEVTEEGTVVQQQPAPAPTALATATAGT-KSVEQEWMTFFSYHTSIW 209
Qy 64 FTVSPNSPGEVLLNLELGPENIPYLAHLARMYNGAGVGVVVLGNAGNAFTAGKIIFAA 123
Db 210 STV---ESQKILYQALNSINPYLDHIAKLYSTWSGGIDVFTVSGSGVGGKLAALL 266
Qy 124 IPPNF-PIDNLSAAQITMCPHVIVDVQLSPVNLPMDFVNNFFHYNQSGDSRLRLIAML 182
Db 267 VPPGVEPESVSMQY---PHVLFDAKQTEPVFTFDIRKTLFHSMDTDT-TKLVINP 322
Qy 183 YTPLRANNSGDDVFTVSCRVLTSPDPSFNFVLP 218
Db 323 Y-----ENGVENKTTCSITVETRFSADETFALLKPP 353

RESULT 12
VCWVFF
coat protein - feline calicivirus (strain CFI/68 FIV)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Aug-1999
C:Accession: A40507; B40507; T09246
R:Neill, J.D.; Reardon, I.M.; Heinrikson, R.L.
J. Virol. 65, 5440-5447, 1991
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline caliciv
A:Reference number: A40507; MUID:91374597; PMID:1716692
A:Accession: A40507
A:Molecule type: genomic RNA
A:Residues: 1-668 <NE1>
A:Cross-references: GB:M32819; NID:g323874; PIDN:AAA42925.1; PID:g323875
A:Accession: B40507
A:Molecule type: protein
A:Residues: 373-379;403-419;481-489;560-566 <NE2>
R:Neill, J.D.
submitted to the EMBL Data Library, April 1998
A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.
A:Reference number: Z16626
A:Accession: T09246
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: genomic RNA
A:Residues: 1-668 <NE3>
A:Cross-references: EMBL:U13992; NID:g3056875; PIDN:AAAL3993.1; PID:g537256
A:Experimental source: strain CFI/68 FIV
C:Superfamily: feline calicivirus coat protein

A:Cross-references: GB:L09718; NID:G305104; PIDN:AAA16485.1; PID:G305105
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein

```
Query Match          8.9%; Score 257; DB 2; Length 668;
Best Local Similarity 23.6%; Pred. No. 2.6e-11;
Matches 72; Conservative 52; Mismatches 99; Indels 82; Gaps 10;

QY 67 SPRNSGEVLLNLELGPEINPYLAHLARMYNGYAGGFVQVVLGNAFTAGKIIFAAIIPP 126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 STSETQKILFKQSLGPLNPLYLHLKLYVAWSGSVEVRFSISGSGVFGKLAIVVPP 238
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 127 NFPIDNLSAAQITMCCHVIVDROLEPNLPHDPVRNFFHYNQSGSDSLRLIAMLYTEL 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 239 G--VDPVQSTSLQYPHVLFDARQVDFVIFSPDLRSTLYHLMPDQDT-TSLVIMVYNDL 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 187 -----RANNSGDDVFTVSCRVLTRPSPDFSFNPLVPP----- 218
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 296 INPYANDSNSSGCIVTVE---TKGPDFKFHLLKPPGSMLTGSGVPSDLIPKSSSLWIG 351

QY 219 -----TVESKTKPFTL-----PI-LTISEMSNSRFPVPPIESLH 250
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 352 NRYWSDITDFVVRPFVQANRHFNFQETAGWSAPRFRPITITISESGSKLGIGVATDY 411

QY 251 -----TSPTENIVVQCONGRVTLGELMGTTQLLPSCAIFRGV-LTR 292
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 412 IVFGIPDGWPDTTIAEDLTAPAGDYAITSGNGNDITTGSEYDSTEVIKNN-TNFRGMYICG 470

QY 293 STSRA 297
   | |
Db 471 SLQRA 475
```

RESULT 15
JQ2356
capsid protein - feline calicivirus (strain KCD)
C:Species: feline calicivirus
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: JQ2356
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.
J:Gen. Virol. 74, 2519-2524, 1993
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable
A:Reference number: JQ2354; MUID:94065683; PMID:7504075
A:Accession: JQ2356
A:Molecule type: mRNA
A:Residues: 1-668 <SEA>
A:Cross-references: GB:L09719; NID:G305107; PIDN:AAA16487.1; PID:G305108
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein

```
Query Match          8.8%; Score 255; DB 2; Length 668;
Best Local Similarity 33.8%; Pred. No. 3.6e-11;
Matches 53; Conservative 31; Mismatches 61; Indels 12; Gaps 4;

QY 67 SPRNSGEVLLNLELGPEINPYLAHLARMYNGYAGGFVQVVLGNAFTAGKIIFAAIIPP 126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 STSETQKILFKQSLGPLNPLYLHLKLYVAWSGSIEVRFSISGSGVFGKLAIVVPP 238
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 127 NFPIDNLSAAQITMCCHVIVDROLEPNLPHDPVRNFFHYNQSGSDSLRLIAMLYTEL 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 239 G--VDPVQSTSLQYPHVLFDARQVDFVIFSPDLRSTLYHLMSDQDT-TSLVVMAYNDL 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 187 -----RANNSGDDVFTVSCRVLTRPSPDFSFNPLVPP 218
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 296 INPYANDSNSSGCIVTVE---TKGSDRFRHLLKPP 328
```

Search completed: June 1, 2004, 13:55:25
Job time : 12.5738 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:35:27 ; Search time 31.3814 Seconds
(without alignments)
5419.275 Million cell updates/sec

Title: US-09-926-799-5
Perfect score: 2886
Sequence: 1 MKMASNDANPSDGSSTANLVP.....VNQFTYLPAMGNGARRRAL 539

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2886	100.0	539	12 Q8JW43	Q8JW43 norwalk-lik
2	2872	99.5	539	12 Q9W9T8	Q9W9T8 human calic
3	2871	99.5	539	12 Q915C5	Q915C5 human calic
4	2871	99.5	539	12 Q9QY55	Q9QY55 human calic
5	2867	99.3	539	12 Q9WR20	Q9WR20 human calic
6	2867	99.3	539	12 Q9W9Y9	Q9W9Y9 human calic
7	2867	99.3	539	12 Q9WR21	Q9WR21 human calic
8	2865	99.3	539	12 Q9WRV9	Q9WRV9 human calic
9	2863	99.2	539	12 Q915C2	Q915C2 human calic
10	2862	99.2	539	12 Q916B8	Q916B8 human calic
11	2860	99.1	539	12 Q916E7	Q916E7 human calic
12	2859	99.1	539	12 Q915D1	Q915D1 human calic
13	2857	99.0	539	12 Q913Z3	Q913Z3 human calic
14	2856	99.0	539	12 Q917W2	Q917W2 norwalk-lik
15	2856	99.0	539	12 Q917W5	Q917W5 norwalk-lik
16	2854	98.9	539	12 Q916F0	Q916F0 human calic

17	2852	98.8	539	12 Q915C3	Q915C3 human calic
18	2852	98.8	539	12 Q916B9	Q916B9 human calic
19	2851	98.8	539	12 Q9WRX8	Q9WRX8 human calic
20	2842	98.5	539	12 Q91V44	Q91V44 human calic
21	2839	98.4	539	12 Q915C9	Q915C9 human calic
22	2838	98.3	539	12 Q9YQSO	Q9YQSO human calic
23	2832	98.1	539	12 Q915D0	Q915D0 human calic
24	2828	98.0	539	12 Q915C4	Q915C4 human calic
25	2828	98.0	539	12 Q91V38	Q91V38 human calic
26	2816	97.6	539	12 Q918V4	Q918V4 norwalk-lik
27	2804	97.2	539	12 Q66287	Q66287 calicivirid
28	2803	97.1	539	12 Q8V0P2	Q8V0P2 human calic
29	2800	97.0	539	12 Q68537	Q68537 human calic
30	2799.5	97.0	540	12 Q915C8	Q915C8 human calic
31	2799	97.0	525	12 Q9QAX3	Q9QAX3 norwalk-lik
32	2799	97.0	539	12 Q8V0P6	Q8V0P6 human calic
33	2770	96.0	539	12 Q917Y6	Q917Y6 norwalk-lik
34	1967.5	68.2	548	12 Q8V768	Q8V768 norwalk vir
35	1963.5	68.0	548	12 Q8V0P5	Q8V0P5 human calic
36	1962.5	68.0	548	12 Q8V0P4	Q8V0P4 human calic
37	1957.5	67.8	548	12 Q66296	Q66296 calicivirid
38	1956.5	67.8	548	12 Q917Z2	Q917Z2 norwalk-lik
39	1954.5	67.7	548	12 Q917Z5	Q917Z5 norwalk-lik
40	1946.5	67.4	548	12 Q8V771	Q8V771 norwalk vir
41	1946.5	67.4	548	12 Q8BC87	Q8BC87 human calic
42	1945.5	67.4	548	12 Q9PYA7	Q9PYA7 human calic
43	1944.5	67.4	548	12 Q91V40	Q91V40 human calic
44	1943.5	67.3	548	12 Q83880	Q83880 norwalk vir
45	1942.5	67.3	548	12 Q916B5	Q916B5 human calic

ALIGNMENTS

RESULT 1

Q8JW43 PRELIMINARY; PRT; 539 AA.
AC Q8JW43
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Capsid.
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
RN [1]_TaxID=95340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Narital04;
RA Katayama K., Takeda N., Natori K.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Narital04;
RA Natori K., Takeda N.;
RT "Genetic and antigenic relationship among Norwalk-like viruses.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB078336; BAC05517.1;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 539 AA; 58815 MW; D9DCE615860F20E9 CRC64;

Query Match 100.0%; Score 2886; DB 12; Length 539;

Best Local Similarity 100.0%; Pred. No. 3.8e-216;

Matches 539; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGSSTANLVPVNVNEVMALEPVVGAIAAPVAGQVQVQVDPWIRNNFVQAP 60

Db 1 MKMASNDANPSDGSSTANLVPVNVNEVMALEPVVGAIAAPVAGQVQVQVDPWIRNNFVQAP 60

QY 61 GGEFTVSPRNAPGEILWSAPGLPDNLNLYSLHARMYVAGGFEVQVILAGNAFTAGKTI 120

Db 61 GGEFTVSPRNAPGEILWSAPLGPDLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120
Qy 121 FFAVPPNPTTEGLSPQVMTFPHIIIVDVRQLEPVLIPLPDVRNNFYHYNQNSDSTIKLIA 180
Db 121 FFAVPPNPTTEGLSPQVMTFPHIIIVDVRQLEPVLIPLPDVRNNFYHYNQNSDSTIKLIA 180
Qy 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFIPLVPPPTVESRTKPTFTVPILTVEEMNS 240
Db 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFIPLVPPPTVESRTKPTFTVPILTVEEMNS 240
Qy 241 RFPPIPLEKLYTGPSSAFVQVQNGRCCTTGVLLGTQTSVAVNICTFRGDVTIAGSHDYT 300
Db 241 RFPPIPLEKLYTGPSSAFVQVQNGRCCTTGVLLGTQTSVAVNICTFRGDVTIAGSHDYT 300
Qy 301 MNLASQWNSNYDPTTEIPAPLGTDFVGVIOQGNHNEPQVWLFQVWQHFCEAAQAQSDVALLRFVNP 360
Db 301 MNLASQWNSNYDPTTEIPAPLGTDFVGVIOQGNHNEPQVWLFQVWQHFCEAAQAQSDVALLRFVNP 360
Qy 361 KLGSVQYTTDTNNDFTQNTKFTPVGVIOQGNHNEPQVWLFQVWQHFCEAAQAQSDVALLRFVNP 420
Db 361 KLGSVQYTTDTNNDFTQNTKFTPVGVIOQGNHNEPQVWLFQVWQHFCEAAQAQSDVALLRFVNP 420
Qy 421 VAPTFEGEQLLFFRSTMPGCGYPNNMLDCLLPQEWQHFCEAAQAQSDVALLRFVNP 480
Db 421 VAPTFEGEQLLFFRSTMPGCGYPNNMLDCLLPQEWQHFCEAAQAQSDVALLRFVNP 480
Qy 481 TGRVLFCEKLYKSGYVTVHTGPHDLVIPPNGYFRFDSWVWQVYTLAPMNGAGRRRAL 539
Db 481 TGRVLFCEKLYKSGYVTVHTGPHDLVIPPNGYFRFDSWVWQVYTLAPMNGAGRRRAL 539

RESULT 2
Q9W9T8 PRELIMINARY; PRT; 539 AA.
AC Q9W9T8; 1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/Mora/97/SE,
OS Human calicivirus isolates, and
OS Human calicivirus NLV/Oberschleissheim/112/99/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=189467, 150080, 173925;
RN [1]
RN SEQUENCE FROM N.A.
RC SPECIES=Human calicivirus NLV/Mora/97/SE; STRAIN=Hu/NLV/Mora/97/SE;
RA Nilsson M., Torken M., Thorhagen M., Hedlund K.-O., Svensson L.;
RT "Expression, self assembly and intracellular location in mammalian cells of a Lordsdale virus-like calicivirus protein.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RC SPECIES=Human calicivirus isolates; STRAIN=Various strains;
RX MEDLINE=99246332; PubMed=10228052;
RA Noel J.S., Fankhauser R.L., Ando I., Monroe S.S., Glass R.I.;
RT "Identification of a distinct common strain of 'Norwalk-like viruses' having a global distribution.";
RL J. Infect. Dis. 179:1334-1344 (1999).
RN [3]
RN SEQUENCE FROM N.A.
RC SPECIES=Human calicivirus NLV/Oberschleissheim/112/99/DE;
RA Kuenkel U., Hoehne M., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF081134; AL87461.1; -
DR EMBL; AF080559; AAD40496.1; -
DR EMBL; AF080549; AAD40488.1; -
DR EMBL; AF080551; AAD40490.1; -

DR EMBL; AF080555; AAD40494.1; -
DR EMBL; AF080557; AAD40496.1; -
DR EMBL; AF427113; AAL19869.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 539 AA; 58902 MW; 0C6CDFA3D6CC345B CRC64;
Query Match 99.5%; Score 2872; DB 12; Length 539;
Best Local Similarity 99.8%; Pred. No. 4.7e-215;
Matches 537; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MKMASNDANPDSGSTANLVPEVNVNVALEPVVGAIAAPVAGQONVIDPWIRNNVQAP 60
Db 1 MKMASNDANPDSGSTANLVPEVNVNVALEPVVGAIAAPVAGQONVIDPWIRNNVQAP 60
Qy 61 GGEFTVSPRNAPGEILWSAPLGPDLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120
Db 61 GGEFTVSPRNAPGEILWSAPLGPDLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120
Qy 121 FFAVPPNPTTEGLSPQVMTFPHIIIVDVRQLEPVLIPLPDVRNNFYHYNQNSDSTIKLIA 180
Db 121 FFAVPPNPTTEGLSPQVMTFPHIIIVDVRQLEPVLIPLPDVRNNFYHYNQNSDSTIKLIA 180
Qy 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFIPLVPPPTVESRTKPTFTVPILTVEEMNS 240
Db 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFIPLVPPPTVESRTKPTFTVPILTVEEMNS 240
Qy 241 RFPPIPLEKLYTGPSSAFVQVQNGRCCTTGVLLGTQTSVAVNICTFRGDVTIAGSHDYT 300
Db 241 RFPPIPLEKLYTGPSSAFVQVQNGRCCTTGVLLGTQTSVAVNICTFRGDVTIAGSHDYT 300
Qy 301 MNLASQWNSNYDPTTEIPAPLGTDFVGVIOQGNHNEPQVWLFQVWQHFCEAAQAQSDVALLRFVNP 360
Db 301 MNLASQWNSNYDPTTEIPAPLGTDFVGVIOQGNHNEPQVWLFQVWQHFCEAAQAQSDVALLRFVNP 360
Qy 361 KLGSVQYTTDTNNDFTQNTKFTPVGVIOQGNHNEPQVWLFQVWQHFCEAAQAQSDVALLRFVNP 420
Db 361 KLGSVQYTTDTNNDFTQNTKFTPVGVIOQGNHNEPQVWLFQVWQHFCEAAQAQSDVALLRFVNP 420
Qy 421 VAPTFEGEQLLFFRSTMPGCGYPNNMLDCLLPQEWQHFCEAAQAQSDVALLRFVNP 480
Db 421 VAPTFEGEQLLFFRSTMPGCGYPNNMLDCLLPQEWQHFCEAAQAQSDVALLRFVNP 480
Qy 481 TGRVLFCEKLYKSGYVTVHTGPHDLVIPPNGYFRFDSWVWQVYTLAPMNGAGRRRAL 539
Db 481 TGRVLFCEKLYKSGYVTVHTGPHDLVIPPNGYFRFDSWVWQVYTLAPMNGAGRRRAL 539

RESULT 3
Q915C5 PRELIMINARY; PRT; 539 AA.
AC Q915C5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/Beeskow/124/00/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=173931;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=NLV/Beeskow/124/00/DE;
RA Kuenkel U., Hoehne M., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF427120; AAL19876.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 539 AA; 58916 MW; CF6B06BE45C17D8F CRC64;

```
Query Match          99.5%; Score 2871; DB 12; Length 539;
Best Local Similarity 99.4%; Pred. No. 5.7e-215;
Matches 536; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGSSTANLVPEVNVNMALEPVVGAIAAPVAGQONVDPWIRNPFVQAP 60
DB 1 MKMASNDANPSDGSSTANLVPEVNVNMALEPVVGAIAAPVAGQONVDPWIRNPFVQAP 60
QY 61 GGEFTVSPRNAPGEILWSAPGLDNLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120
DB 61 GGEFTVSPRNAPGEILWSAPGLDNLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120
QY 121 FFAVPPNPFTEGLSPSQVTMFPHIIVDVROLEPVLIPLDVVRNPFYHYNQSNDSIKLIA 180
DB 121 FFAVPPNPFTEGLSPSQVTMFPHIIVDVROLEPVLIPLDVVRNPFYHYNQSNDSIKLIA 180
QY 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFILVPPPTVESRTKPTFVPIILTVEMSN 240
DB 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFILVPPPTVESRTKPTFVPIILTVEMSN 240
QY 241 RPPLEKLYTGPSSAFVVPQNGRCTTGDVLLGTTQLSAVNICFRGDTVTHIAGSHDYT 300
DB 241 RPPLEKLYTGPSSAFVVPQNGRCTTGDVLLGTTQLSAVNICFRGDTVTHIAGSHDYT 300
QY 301 MNLASQNSNYDPTBEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGSHVFT 360
DB 301 MNLASQNSNYDPTBEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGSHVFT 360
QY 361 KLGSVQYTTDTNDFQGTQNTKFTPVGVIQDGNHNEPQOQVLPNYSRGTGHNVHLAPA 420
DB 361 KLGSVQYTTDTNDFQGTQNTKFTPVGVIQDGNHNEPQOQVLPNYSRGTGHNVHLAPA 420
QY 421 VAPTPPGEQLLFFRSTMPGCSGYPNNMLDCLLPQEWVQHFQCEAAAPQSDVALLRFVND 480
DB 421 VAPTPPGEQLLFFRSTMPGCSGYPNNMLDCLLPQEWVQHFQCEAAAPQSDVALLRFVND 480
QY 481 TGRVLFECKLHKSGYVTVVHTGPHDLVIPPNGYFRFDSVWVQFYTLAPMNGAGRRRAL 539
DB 481 TGRVLFECKLHKSGYVTVVHTGPHDLVIPPNGYFRFDSVWVQFYTLAPMNGAGRRRAL 539

RESULT 4
QYQY55 ID Q8QY55 PRELIMINARY; PRT; 539 AA.
AC Q8QY55;
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/DJON171/96.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=186843;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/DJON171/96;
RA Nicollier-Damot B., Pico V., Pothier P., Kohli B.;
RT "Molecular cloning, expression, self-assembly and seroepidemiology of a genogroup II Norwalk-like virus in France.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF472623; AAL79839.1;
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 539 AA; 58916 MW; 801D18BFA315CAF4 CRC64;

Query Match          99.5%; Score 2871; DB 12; Length 539;
Best Local Similarity 99.4%; Pred. No. 5.7e-215;
Matches 536; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGSSTANLVPEVNVNMALEPVVGAIAAPVAGQONVDPWIRNPFVQAP 60
DB 1 MKMASNDANPSDGSSTANLVPEVNVNMALEPVVGAIAAPVAGQONVDPWIRNPFVQAP 60
QY 61 GGEFTVSPRNAPGEILWSAPGLDNLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120
DB 61 GGEFTVSPRNAPGEILWSAPGLDNLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120
QY 121 FFAVPPNPFTEGLSPSQVTMFPHIIVDVROLEPVLIPLDVVRNPFYHYNQSNDSIKLIA 180
DB 121 FFAVPPNPFTEGLSPSQVTMFPHIIVDVROLEPVLIPLDVVRNPFYHYNQSNDSIKLIA 180
QY 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFILVPPPTVESRTKPTFVPIILTVEMSN 240
DB 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFILVPPPTVESRTKPTFVPIILTVEMSN 240
QY 241 RPPLEKLYTGPSSAFVVPQNGRCTTGDVLLGTTQLSAVNICFRGDTVTHIAGSHDYT 300
DB 241 RPPLEKLYTGPSSAFVVPQNGRCTTGDVLLGTTQLSAVNICFRGDTVTHIAGSHDYT 300
QY 301 MNLASQNSNYDPTBEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGSHVFT 360
DB 301 MNLASQNSNYDPTBEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGSHVFT 360
QY 361 KLGSVQYTTDTNDFQGTQNTKFTPVGVIQDGNHNEPQOQVLPNYSRGTGHNVHLAPA 420
DB 361 KLGSVQYTTDTNDFQGTQNTKFTPVGVIQDGNHNEPQOQVLPNYSRGTGHNVHLAPA 420
QY 421 VAPTPPGEQLLFFRSTMPGCSGYPNNMLDCLLPQEWVQHFQCEAAAPQSDVALLRFVND 480
DB 421 VAPTPPGEQLLFFRSTMPGCSGYPNNMLDCLLPQEWVQHFQCEAAAPQSDVALLRFVND 480
QY 481 TGRVLFECKLHKSGYVTVVHTGPHDLVIPPNGYFRFDSVWVQFYTLAPMNGAGRRRAL 539
DB 481 TGRVLFECKLHKSGYVTVVHTGPHDLVIPPNGYFRFDSVWVQFYTLAPMNGAGRRRAL 539

RESULT 5
QYQY55 ID Q9WR20 PRELIMINARY; PRT; 539 AA.
AC Q9WR20;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus isolates.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=150080;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=358/96015107/1996/FL;
RX MEDLINE=9246332; PubMed=10228052;
RA Noel H.F.S., Frankhauser R., L'Amour J., Montroies S., Goggin R.,
RT "Identification of a novel and distinct common Norwalk-like virus,
RL U.S. Infect. Dis. 279:3339-3343 (1999).
DR EMBL; AF080552; AAD40491.1;
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 539 AA; 58875 MW; 0C6CDA3CA1DF95B CRC64;

Query Match          99.3%; Score 2867; DB 12; Length 539;
Best Local Similarity 99.4%; Pred. No. 1.2e-214;
Matches 536; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGSSTANLVPEVNVNMALEPVVGAIAAPVAGQONVDPWIRNPFVQAP 60
DB 1 MKMASNDANPSDGSSTANLVPEVNVNMALEPVVGAIAAPVAGQONVDPWIRNPFVQAP 60
QY 61 GGEFTVSPRNAPGEILWSAPGLDNLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120
DB 61 GGEFTVSPRNAPGEILWSAPGLDNLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120
```

QY	121	FAAVPPNPFTEGLSPSQVTMPFHHIIVDVRQLEPVLIPLEDVVRNNFYHYNQSDSTIKLIA	180
Db	121	FAAVPPNPFTEGLSPSQVTMPFHHIIVDVRQLEPVLIPLEDVVRNNFYHYNQSDSTIKLIA	180
QY	181	MLYTPLRANNAGDDVFTVSCRVLTRPSDPDFIFLVPPTVESRTKPTFTVPILTVEEMSN	240
Db	181	MLYTPLRANNAGDDVFTVSCRVLTRPSDPDFIFLVPPTVESRTKPTFTVPILTVEEMSN	240
QY	241	RPFIPEKLYTGPSAFVVPQNGRCTTGVLLGTTQLSAVNICFRGDVTHIAGSHDYT	300
Db	241	RPFIPEKLYTGPSAFVVPQNGRCTTGVLLGTTQLSAVNICFRGDVTHIAGSHDYT	300
QY	301	MNLASQWNSNYDPTBEIPAPLGTDFVGVKIQQMLTQTTREDGSTRAHKATVSTGSHVFTP	360
Db	301	MNLASQWNSNYDPTBEIPAPLGTDFVGVKIQQMLTQTTREDGSTRAHKATVSTGSHVFTP	360
QY	361	KLGSVQYTTDTNNDPOTGQNTKFTPVGVIOGNNHNEPQQWVLPNYSRGTGHNVLHAPA	420
Db	361	KLGSVQYTTDTNNDPOTGQNTKFTPVGVIOGNNHNEPQQWVLPNYSRGTGHNVLHAPA	420
QY	421	VAPTFGEQLLFRSTMPGCGYPNNMLDCLLPQEWVQHFQCEAAAPQASDVALLRFVNDP	480
Db	421	VAPTFGEQLLFRSTMPGCGYPNNMLDCLLPQEWVQHFQCEAAAPQASDVALLRFVNDP	480
QY	481	TGRVLFECKLHKSQYVTVTAHTGPHDLVIPPNGYFRFDSWVNWQYITLAPMNGAGRRAL	539
Db	481	TGRVLFECKLHKSQYVTVTAHTGPHDLVIPPNGYFRFDSWVNWQYITLAPMNGAGRRAL	539
RESULT 6			
ID	Q9W9Y9	PRELIMINARY; PRT; 539 AA.	
AC	Q9W9Y9	01-NOV-1999 (TREMBLrel. 12, Created)	
DT	01-NOV-1999	(TREMBLrel. 12, Last sequence update)	
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)	
DE	Capsid protein.		
OS	Human calicivirus isolates.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;		
OC	Norovirus.		
OX	NCBI_taxid=150080;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=379/96019984/1996/AZ, and 364/96019537/1996/AZ;		
RX	MEDLINE=93246332; PubMed=10228052;		
RA	Noel J.S., Fankhauser R.L., Ando T., Monroe S.S., Glass R.I.;		
RT	"Identification of a distinct common strain of 'Norwalk-like viruses' having a global distribution."		
RL	J. Infect. Dis. 179:1334-1344 (1999).		
DR	EMBL; AF080556; AAD40495.1; -.		
DR	EMBL; AF080553; AAD40492.1; -.		
DR	InterPro; IPR004005; Calici_coat.		
DR	InterPro; IPR008975; Viral_Cap_coat.		
DR	Pfam; PF00915; Calici_coat; 1.		
SQ	SEQUENCE 539 AA; 58872 MW; 1D6CD304101BEF2B CRC64;		
Query Match 99.3%; Score 2867; DB 12; Length 539;			
Best Local Similarity 99.4%; Pred. No. 1.2e-214;			
Matches 536; Conservative 1; Mismatches 2; Indels 0; Gaps 0;			
QY	1	MKMASNDANPSDGSSTANLVPEVNNVMALEPVVGAIAAPVAGQONVIDPWRNFVQAP	60
Db	1	MKMASNDANPSDGSSTANLVPEVNNVMALEPVVGAIAAPVAGQONVIDPWRNFVQAP	60
QY	61	GGEFTVSPRNAPGEILWSAPLGPDLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII	120
Db	61	GGEFTVSPRNAPGEILWSAPLGPDLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII	120
QY	121	FAAVPPNPFTEGLSPSQVTMPFHHIIVDVRQLEPVLIPLEDVVRNNFYHYNQSDSTIKLIA	180
Db	121	FAAVPPNPFTEGLSPSQVTMPFHHIIVDVRQLEPVLIPLEDVVRNNFYHYNQSDSTIKLIA	180
QY	181	MLYTPLRANNAGDDVFTVSCRVLTRPSDPDFIFLVPPTVESRTKPTFTVPILTVEEMSN	240
Db	181	MLYTPLRANNAGDDVFTVSCRVLTRPSDPDFIFLVPPTVESRTKPTFTVPILTVEEMSN	240
QY	241	RPFIPEKLYTGPSAFVVPQNGRCTTGVLLGTTQLSAVNICFRGDVTHIAGSHDYT	300
Db	241	RPFIPEKLYTGPSAFVVPQNGRCTTGVLLGTTQLSAVNICFRGDVTHIAGSHDYT	300

Db	181	MLYTPLRANNAGDDVFTVSCRVLTRPSDPDFIFLVPPTVESRTKPTFTVPILTVEEMSN	240
QY	241	RPFIPEKLYTGPSAFVVPQNGRCTTGVLLGTTQLSAVNICFRGDVTHIAGSHDYT	300
Db	241	RPFIPEKLYTGPSAFVVPQNGRCTTGVLLGTTQLSAVNICFRGDVTHIAGSHDYT	300
QY	301	MNLASQWNSNYDPTBEIPAPLGTDFVGVKIQQMLTQTTREDGSTRAHKATVSTGSHVFTP	360
Db	301	MNLASQWNSNYDPTBEIPAPLGTDFVGVKIQQMLTQTTREDGSTRAHKATVSTGSHVFTP	360
QY	361	KLGSVQYTTDTNNDPOTGQNTKFTPVGVIOGNNHNEPQQWVLPNYSRGTGHNVLHAPA	420
Db	361	KLGSVQYTTDTNNDPOTGQNTKFTPVGVIOGNNHNEPQQWVLPNYSRGTGHNVLHAPA	420
QY	421	VAPTFGEQLLFRSTMPGCGYPNNMLDCLLPQEWVQHFQCEAAAPQASDVALLRFVNDP	480
Db	421	VAPTFGEQLLFRSTMPGCGYPNNMLDCLLPQEWVQHFQCEAAAPQASDVALLRFVNDP	480
QY	481	TGRVLFECKLHKSQYVTVTAHTGPHDLVIPPNGYFRFDSWVNWQYITLAPMNGAGRRAL	539
Db	481	TGRVLFECKLHKSQYVTVTAHTGPHDLVIPPNGYFRFDSWVNWQYITLAPMNGAGRRAL	539
RESULT 7			
ID	Q9WRZ1	PRELIMINARY; PRT; 539 AA.	
AC	Q9WRZ1	01-NOV-1999 (TREMBLrel. 12, Created)	
DT	01-NOV-1999	(TREMBLrel. 12, Last sequence update)	
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)	
DE	Capsid protein.		
OS	Human calicivirus isolates.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;		
OC	Norovirus.		
OX	NCBI_taxid=150080;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=345-2/96002737/1996/SC;		
RX	MEDLINE=99246332; PubMed=10228052;		
RA	Noel J.S., Fankhauser R.L., Ando T., Monroe S.S., Glass R.I.;		
RT	"Identification of a distinct common strain of 'Norwalk-like viruses' having a global distribution."		
RL	J. Infect. Dis. 179:1334-1344 (1999).		
DR	EMBL; AF080550; AAD40489.1; -.		
DR	InterPro; IPR004005; Calici_coat.		
DR	InterPro; IPR008975; Viral_Cap_coat.		
DR	Pfam; PF00915; Calici_coat; 1.		
SQ	SEQUENCE 539 AA; 58875 MW; 072C697DD93F3BA7 CRC64;		
Query Match 99.3%; Score 2867; DB 12; Length 539;			
Best Local Similarity 99.4%; Pred. No. 1.2e-214;			
Matches 536; Conservative 2; Mismatches 1; Indels 0; Gaps 0;			
QY	1	MKMASNDANPSDGSSTANLVPEVNNVMALEPVVGAIAAPVAGQONVIDPWRNFVQAP	60
Db	1	MKMASNDANPSDGSSTANLVPEVNNVMALEPVVGAIAAPVAGQONVIDPWRNFVQAP	60
QY	61	GGEFTVSPRNAPGEILWSAPLGPDLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII	120
Db	61	GGEFTVSPRNAPGEILWSAPLGPDLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII	120
QY	121	FAAVPPNPFTEGLSPSQVTMPFHHIIVDVRQLEPVLIPLEDVVRNNFYHYNQSDSTIKLIA	180
Db	121	FAAVPPNPFTEGLSPSQVTMPFHHIIVDVRQLEPVLIPLEDVVRNNFYHYNQSDSTIKLIA	180
QY	181	MLYTPLRANNAGDDVFTVSCRVLTRPSDPDFIFLVPPTVESRTKPTFTVPILTVEEMSN	240
Db	181	MLYTPLRANNAGDDVFTVSCRVLTRPSDPDFIFLVPPTVESRTKPTFTVPILTVEEMSN	240
QY	241	RPFIPEKLYTGPSAFVVPQNGRCTTGVLLGTTQLSAVNICFRGDVTHIAGSHDYT	300
Db	241	RPFIPEKLYTGPSAFVVPQNGRCTTGVLLGTTQLSAVNICFRGDVTHIAGSHDYT	300

```

QY 301 MNLASQNSNYDPTBEIPAPLGTDPFVGKI QGMLTQTTREDGSTRAHKATVSTGSHVFTT 360
DB 301 MNLASQNSNYDPTBEIPAPLGTDPFVGKI QGMLTQTTREDGSTRAHKATVSTGSHVFTT 360
QY 361 KLGSGVQYTTDTNNDFTQNTKFTPVGVI QDGNHNEPQQWVLPNYSGRTHNVHLAPA 420
DB 361 KLGSGVQYTTDTNNDFTQNTKFTPVGVI QDGNHNEPQQWVLPNYSGRTHNVHLAPA 420
QY 421 VAPTFPGEQLLFFRSTMPGCGSGYPNNLDCLLLPQEWVQHFCEAAAPQSDVALLRFVNDP 480
DB 421 VAPTFPGEQLLFFRSTMPGCGSGYPNNLDCLLLPQEWVQHFCEAAAPQSDVALLRFVNDP 480
QY 481 TGRVLFECKLHKSQYVTVHTGPHDLVTPPNQYFRFDSWVNFYTLAPMNGAGRRRAL 539
DB 481 TGRVLFECKLHKSQYVTVHTGPHDLVTPPNQYFRFDSWVNFYTLAPMNGAGRRRAL 539

RESULT 8
QWRY9
ID QWRY9 PRELIMINARY; PRT; 539 AA.
AC QWRY9;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus isolates.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=150080;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=366/96019554/1996/ID;
RX MEDLINE=99246332; PubMed=10228052;
RA Noel J.S., Fankhauser R.L., Ando T., Monroe S.S., Glass R.I.;
RT "Identification of a distinct common strain of 'Norwalk-like viruses'
   having a global distribution.";
RL J. Infect. Dis. 179:1334-1344 (1999).
RE EMBL; AF080554; AAD40493.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat_1;
SQ SEQUENCE 539 AA; 58842 MW; C7B31E3126A8F4FF CRC64;

Query Match 99.3%; Score 2865; DB 12; Length 539;
Best Local Similarity 99.3%; Pred. No. 1.7e-214;
Matches 535; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGSSTANLVPEVNEVMALEPVVGAIAAPVAGQQNVDPWIRNNFVQAP 60
DB 1 MKMASNDANPSDGSSTANLVPEVNEVMALEPVVGAIAAPVAGQQNVDPWIRNNFVQAP 60
QY 61 GGEFTVSPRNAPGELTWSAPLGPDLNPLYSLHARMYNGVAGFEVQVILAGNAFTAGKII 120
DB 61 GGEFTVSPRNAPGELTWSAPLGPDLNPLYSLHARMYNGVAGFEVQVILAGNAFTAGKII 120
QY 121 FFAVPPNPTFEGSLSPSQTMTFPHIIVDVRLPEPLIPLDVRNNFYHYNQNSDSTIKLIA 180
DB 121 FFAVPPNPTFEGSLSPSQTMTFPHIIVDVRLPEPLIPLDVRNNFYHYNQNSDSTIKLIA 180
QY 181 MLYPLRANAGDDVFTVSCRVLTPSPDFDFILVPPVTESRTKPTFTVILTVSEMSNS 240
DB 181 MLYPLRANAGDDVFTVSCRVLTPSPDFDFILVPPVTESRTKPTFTVILTVSEMSNS 240
QY 241 RPIPLEKLYTGPSSAFVVPQNGRCTTGDVLLGTTLQTSAVNICITFRGDVTHIASHDYT 300
DB 241 RPIPLEKLYTGPSSAFVVPQNGRCTTGDVLLGTTLQTSAVNICITFRGDVTHIASHDYT 300
QY 301 MNLASQNSNYDPTBEIPAPLGTDPFVGKI QGMLTQTTREDGSTRAHKATVSTGSHVFTT 360
DB 301 MNLASQNSNYDPTBEIPAPLGTDPFVGKI QGMLTQTTREDGSTRAHKATVSTGSHVFTT 360
QY 361 KLGSGVQYTTDTNNDFTQNTKFTPVGVI QDGNHNEPQQWVLPNYSGRTHNVHLAPA 420
DB 361 KLGSGVQYTTDTNNDFTQNTKFTPVGVI QDGNHNEPQQWVLPNYSGRTHNVHLAPA 420

```

```

DB 361 KLGSGVQYTTDTNNDFTQNTKFTPVGVI QDGNHNEPQQWVLPNYSGRTHNVHLAPA 420
QY 421 VAPTFPGEQLLFFRSTMPGCGSGYPNNLDCLLLPQEWVQHFCEAAAPQSDVALLRFVNDP 480
DB 421 VAPTFPGEQLLFFRSTMPGCGSGYPNNLDCLLLPQEWVQHFCEAAAPQSDVALLRFVNDP 480
QY 481 TGRVLFECKLHKSQYVTVHTGPHDLVTPPNQYFRFDSWVNFYTLAPMNGAGRRRAL 539
DB 481 TGRVLFECKLHKSQYVTVHTGPHDLVTPPNQYFRFDSWVNFYTLAPMNGAGRRRAL 539

RESULT 9
QWRY9
ID QWRY9 PRELIMINARY; PRT; 539 AA.
AC QWRY9;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/Berlin/495/00/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=173934;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/Berlin/495/00/DE;
RA Kuenkel U., Hoehne M., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated
   with Norwalk-like viruses in Germany.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBSJ databases.
RE EMBL; AF427123; AAL18879.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat_1;
SQ SEQUENCE 539 AA; 58869 MW; BC145CA144699DAC CRC64;

Query Match 99.2%; Score 2863; DB 12; Length 539;
Best Local Similarity 99.3%; Pred. No. 2.4e-214;
Matches 535; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGSSTANLVPEVNEVMALEPVVGAIAAPVAGQQNVDPWIRNNFVQAP 60
DB 1 MKMASNDANPSDGSSTANLVPEVNEVMALEPVVGAIAAPVAGQQNVDPWIRNNFVQAP 60
QY 61 GGEFTVSPRNAPGELTWSAPLGPDLNPLYSLHARMYNGVAGFEVQVILAGNAFTAGKII 120
DB 61 GGEFTVSPRNAPGELTWSAPLGPDLNPLYSLHARMYNGVAGFEVQVILAGNAFTAGKII 120
QY 121 FFAVPPNPTFEGSLSPSQTMTFPHIIVDVRLPEPLIPLDVRNNFYHYNQNSDSTIKLIA 180
DB 121 FFAVPPNPTFEGSLSPSQTMTFPHIIVDVRLPEPLIPLDVRNNFYHYNQNSDSTIKLIA 180
QY 181 MLYPLRANAGDDVFTVSCRVLTPSPDFDFILVPPVTESRTKPTFTVILTVSEMSNS 240
DB 181 MLYPLRANAGDDVFTVSCRVLTPSPDFDFILVPPVTESRTKPTFTVILTVSEMSNS 240
QY 241 RPIPLEKLYTGPSSAFVVPQNGRCTTGDVLLGTTLQTSAVNICITFRGDVTHIASHDYT 300
DB 241 RPIPLEKLYTGPSSAFVVPQNGRCTTGDVLLGTTLQTSAVNICITFRGDVTHIASHDYT 300
QY 301 MNLASQNSNYDPTBEIPAPLGTDPFVGKI QGMLTQTTREDGSTRAHKATVSTGSHVFTT 360
DB 301 MNLASQNSNYDPTBEIPAPLGTDPFVGKI QGMLTQTTREDGSTRAHKATVSTGSHVFTT 360
QY 361 KLGSGVQYTTDTNNDFTQNTKFTPVGVI QDGNHNEPQQWVLPNYSGRTHNVHLAPA 420
DB 361 KLGSGVQYTTDTNNDFTQNTKFTPVGVI QDGNHNEPQQWVLPNYSGRTHNVHLAPA 420
QY 421 VAPTFPGEQLLFFRSTMPGCGSGYPNNLDCLLLPQEWVQHFCEAAAPQSDVALLRFVNDP 480
DB 421 VAPTFPGEQLLFFRSTMPGCGSGYPNNLDCLLLPQEWVQHFCEAAAPQSDVALLRFVNDP 480

```

QY 481 TGRVLFPECKLHKSGVYVVAHTGPHDLVIPPNGYFRFDSWVNFQFYTLAPMGNGAGRRAL 539
 Db 481 TGRVLFPECKLHKSGVYVVAHTGPHDLVIPPNGYFRFDSWVNFQFYTLAPMGNGAGRRAL 539

RESULT 10

QY16E8 PRELIMINARY; PRT; 539 AA.

AC Q916E8; (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Capsid protein.
 OS Human calicivirus NLV/Altenkirchen 140/01/DE.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 OC NCBI_TaxID=173918;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NLV/Altenkirchen 140/01/DE;
 RA Kuenkel U., Hoehne M., Schreier E.;
 RT "Molecular epidemiology of outbreaks of gastroenteritis associated
 with Norwalk-like viruses in Germany"; J. Virol. 74:1185-1191 (2000).
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF425765; AAL18858.1; -;
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral Cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 SQ SEQUENCE 539 AA; 5829 MW; 06A596B3A7D83B34 CRC64;

Query Match 99.2%; Score 2862; DB 12; Length 539;
 Best Local Similarity 99.1%; Pred. No. 2.8e-214;
 Matches 534; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGSSTANLVEPVNEVMALPVPVGAIAAPVAGQNVDPWIRNFFVQAP 60
 Db 1 MKMASNDANPSDGSSTANLVEPVNEVMALPVPVGAIAAPVAGQNVDPWIRNFFVQAP 60

QY 61 GGEFTVSPRNAPGELIWSAPLGNPLNPLSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
 Db 61 GGEFTVSPRNAPGELIWSAPLGNPLNPLSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120

QY 121 FAAPVPPNPTTEGLSPSQVTPPHIIVDVROLEPVLIPLDVRNFFHYNSNDSTIKLIA 180
 Db 121 FAAPVPPNPTTEGLSPSQVTPPHIIVDVROLEPVLIPLDVRNFFHYNSNDSTIKLIA 180

QY 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFILVPPTVESRTKPTFVPIITVEEMNS 240
 Db 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFILVPPTVESRTKPTFVPIITVEEMNS 240

QY 241 RPIPLEKLYTGPSAFVQVQNGRCITTDGVLGTTQLSAVNICTFRGDVTHIAGSHDYT 300
 Db 241 RPIPLEKLYTGPSAFVQVQNGRCITTDGVLGTTQLSAVNICTFRGDVTHIAGSHDYT 300

QY 301 MNLASQNSVNDPTEIEIPALGTDPFVGKIQGLMTQTTREDGSTRAHKATVSTGSHVFTF 360
 Db 301 MNLASQNSVNDPTEIEIPALGTDPFVGKIQGLMTQTTREDGSTRAHKATVSTGSHVFTF 360

QY 361 KLGSVQYTTDNDTQNTKFTPVGVIQDGNHNEPQWVLPNTVSGRTGHNHVLAPA 420
 Db 361 KLGSVQYTTDNDTQNTKFTPVGVIQDGNHNEPQWVLPNTVSGRTGHNHVLAPA 420

QY 421 VAPTFPGQLLFFRSTMGCSGYPMNLDCLLPQEWVQHFCQEAAPQSDVALLRFVNP 480
 Db 421 VAPTFPGQLLFFRSTMGCSGYPMNLDCLLPQEWVQHFCQEAAPQSDVALLRFVNP 480

QY 481 TGRVLFPECKLHKSGVYVVAHTGPHDLVIPPNGYFRFDSWVNFQFYTLAPMGNGAGRRAL 539
 Db 481 TGRVLFPECKLHKSGVYVVAHTGPHDLVIPPNGYFRFDSWVNFQFYTLAPMGNGAGRRAL 539

QY16E7 PRELIMINARY; PRT; 539 AA.

AC Q916E7; (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Capsid protein.
 OS Human calicivirus NLV/Dillingen 259/01/DE.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 OC NCBI_TaxID=173919;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NLV/Dillingen 259/01/DE;
 RA Kuenkel U., Hoehne M., Schreier E.;
 RT "Molecular epidemiology of outbreaks of gastroenteritis associated
 with Norwalk-like viruses in Germany"; J. Virol. 74:1185-1191 (2000).
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF425766; AAL18859.1; -;
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral Cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 SQ SEQUENCE 539 AA; 5855 MW; DC1EC814C684E131 CRC64;

Query Match 99.1%; Score 2860; DB 12; Length 539;
 Best Local Similarity 99.1%; Pred. No. 4.1e-214;
 Matches 534; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGSSTANLVEPVNEVMALPVPVGAIAAPVAGQNVDPWIRNFFVQAP 60
 Db 1 MKMASNDANPSDGSSTANLVEPVNEVMALPVPVGAIAAPVAGQNVDPWIRNFFVQAP 60

QY 61 GGEFTVSPRNAPGELIWSAPLGNPLNPLSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
 Db 61 GGEFTVSPRNAPGELIWSAPLGNPLNPLSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120

QY 121 FAAPVPPNPTTEGLSPSQVTPPHIIVDVROLEPVLIPLDVRNFFHYNSNDSTIKLIA 180
 Db 121 FAAPVPPNPTTEGLSPSQVTPPHIIVDVROLEPVLIPLDVRNFFHYNSNDSTIKLIA 180

QY 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFILVPPTVESRTKPTFVPIITVEEMNS 240
 Db 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFILVPPTVESRTKPTFVPIITVEEMNS 240

QY 241 RPIPLEKLYTGPSAFVQVQNGRCITTDGVLGTTQLSAVNICTFRGDVTHIAGSHDYT 300
 Db 241 RPIPLEKLYTGPSAFVQVQNGRCITTDGVLGTTQLSAVNICTFRGDVTHIAGSHDYT 300

QY 301 MNLASQNSVNDPTEIEIPALGTDPFVGKIQGLMTQTTREDGSTRAHKATVSTGSHVFTF 360
 Db 301 MNLASQNSVNDPTEIEIPALGTDPFVGKIQGLMTQTTREDGSTRAHKATVSTGSHVFTF 360

QY 361 KLGSVQYTTDNDTQNTKFTPVGVIQDGNHNEPQWVLPNTVSGRTGHNHVLAPA 420
 Db 361 KLGSVQYTTDNDTQNTKFTPVGVIQDGNHNEPQWVLPNTVSGRTGHNHVLAPA 420

QY 421 VAPTFPGQLLFFRSTMGCSGYPMNLDCLLPQEWVQHFCQEAAPQSDVALLRFVNP 480
 Db 421 VAPTFPGQLLFFRSTMGCSGYPMNLDCLLPQEWVQHFCQEAAPQSDVALLRFVNP 480

QY 481 TGRVLFPECKLHKSGVYVVAHTGPHDLVIPPNGYFRFDSWVNFQFYTLAPMGNGAGRRAL 539
 Db 481 TGRVLFPECKLHKSGVYVVAHTGPHDLVIPPNGYFRFDSWVNFQFYTLAPMGNGAGRRAL 539

RESULT 12
 Q915D1 PRELIMINARY; PRT; 539 AA.

AC Q915D1; (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Capsid protein.
 OS Human calicivirus NLV/Frankfurt (Oder)/170/99/DE.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC Norovirus.

OX NCBI_TaxID=173926;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NLV/Frankfurt (Oder)/170/99/DE;

RA Kuenkel U., Hoehne M., Schreier E.;

RT "Molecular epidemiology of outbreaks of gastroenteritis associated

with Norwalk-like viruses in Germany";

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF427114; AAL18870.1; -;

DR InterPro; IPR004005; Calici_coat.

DR InterPro; IPR008975; Viral_Cap_coat.

DR Pfam; PF00915; Calici_coat; 1.

SQ SEQUENCE 539 AA; 58815 MW; 09E2BD6C2B984090 CRC64;

Query Match 99.1%; Score 2859; DB 12; Length 539;

Best Local Similarity 99.1%; Pred No. 4.9e-214;

Matches 534; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGTANLVPEVNEVMALPEVVGAAIAAPVAGQQNVDPWIRNNFVQAP 60

Db 1 MKMASNDANPSDGTANLVPEVNEVMALPEVVGAAIAAPVAGQQNVDPWIRNNFVQAP 60

QY 61 GGEFTVSPRNAPGELWSAPLGPDLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120

Db 61 GGEFTVSPRNAPGELWSAPLGPDLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120

QY 121 FAAVPPNPTEGLSPSQVTMPFPHIIVDVRLQLEPVLIPLDVRRNFFHYNQSNDSITKIIA 180

Db 121 FAAVPPNPTEGLSPSQVTMPFPHIIVDVRLQLEPVLIPLDVRRNFFHYNQSNDSITKIIA 180

QY 181 MLYTLRANAGDDVFTVSCRVLTRPSDFDFILVPPPTVESRTKPTFPVILTVEMSN 240

Db 181 MLYTLRANAGDDVFTVSCRVLTRPSDFDFILVPPPTVESRTKPTFPVILTVEMSN 240

QY 241 RPPIPLEKLYTGPPSAFVVPQNGRCTTGVLLGTTLQTSVAVNICFRGDVTHIAGSHDYT 300

Db 241 RPPIPLEKLYTGPPSAFVVPQNGRCTTGVLLGTTLQTSVAVNICFRGDVTHIAGSHDYT 300

QY 301 NNLASQNNNDYDPTBEIPAPLGTDPFVGKIQGLTQTTREDGSTRAHKATVSTGSHVFT 360

Db 301 NNLASQNNNDYDPTBEIPAPLGTDPFVGKIQGLTQTTREDGSTRAHKATVSTGSHVFT 360

QY 361 KLGVSQYVYTTDNDTQNTKFTPVGVIQGNHNEPQQWVLPNYSGRTHNVHLAPA 420

Db 361 KLGVSQYVYTTDNDTQNTKFTPVGVIQGNHNEPQQWVLPNYSGRTHNVHLAPA 420

QY 421 VAPTPGQQLLFFRSTMPGCGYPNNLDCLLPQEWVQHFQEAAPASQDVALLRFVNP 480

Db 421 VAPTPGQQLLFFRSTMPGCGYPNNLDCLLPQEWVQHFQEAAPASQDVALLRFVNP 480

QY 481 TGRVLFECKLHKSGYVTVVHTGPHDLVIPPNGYFRFDSWVNOFYTLAPMNGAGRRAL 539

Db 481 TGRVLFECKLHKSGYVTVVHTGPHDLVIPPNGYFRFDSWVNOFYTLAPMNGAGRRAL 539

RESULT 13

Q91323

ID Q91323

AC Q91323 PRELIMINARY; PRT; 539 AA.

DT 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Capsid.

OS Human calicivirus NLV/VA98387/1998.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC Norovirus.

OX NCBI_TaxID=165511;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NLV/VA98387/1998;

RA Jiang X., Zhong W.M., Farkas T., Huang P.W., Wilton N., Barrett E.,

RA Fulton D., Morrow R., Matson D.O.;

RT "Baculovirus expression and antigenic characterization of the capsid

proteins of three Norwalk-like viruses.";

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AV038600; AAK84679.1; -;

DR InterPro; IPR004005; Calici_coat.

DR InterPro; IPR008975; Viral_Cap_coat.

DR Pfam; PF00915; Calici_coat; 1.

SQ SEQUENCE 539 AA; 58887 MW; E9A784B3604C83AC CRC64;

Query Match 99.0%; Score 2857; DB 12; Length 539;

Best Local Similarity 98.9%; Pred No. 7e-214;

Matches 533; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGTANLVPEVNEVMALPEVVGAAIAAPVAGQQNVDPWIRNNFVQAP 60

Db 1 MKMASNDANPSDGTANLVPEVNEVMALPEVVGAAIAAPVAGQQNVDPWIRNNFVQAP 60

QY 61 GGEFTVSPRNAPGELWSAPLGPDLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120

Db 61 GGEFTVSPRNAPGELWSAPLGPDLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120

QY 121 FAAVPPNPTEGLSPSQVTMPFPHIIVDVRLQLEPVLIPLDVRRNFFHYNQSNDSITKIIA 180

Db 121 FAAVPPNPTEGLSPSQVTMPFPHIIVDVRLQLEPVLIPLDVRRNFFHYNQSNDSITKIIA 180

QY 181 MLYTLRANAGDDVFTVSCRVLTRPSDFDFILVPPPTVESRTKPTFPVILTVEMSN 240

Db 181 MLYTLRANAGDDVFTVSCRVLTRPSDFDFILVPPPTVESRTKPTFPVILTVEMSN 240

QY 241 RPPIPLEKLYTGPPSAFVVPQNGRCTTGVLLGTTLQTSVAVNICFRGDVTHIAGSHDYT 300

Db 241 RPPIPLEKLYTGPPSAFVVPQNGRCTTGVLLGTTLQTSVAVNICFRGDVTHIAGSHDYT 300

QY 301 NNLASQNNNDYDPTBEIPAPLGTDPFVGKIQGLTQTTREDGSTRAHKATVSTGSHVFT 360

Db 301 NNLASQNNNDYDPTBEIPAPLGTDPFVGKIQGLTQTTREDGSTRAHKATVSTGSHVFT 360

QY 361 KLGVSQYVYTTDNDTQNTKFTPVGVIQGNHNEPQQWVLPNYSGRTHNVHLAPA 420

Db 361 KLGVSQYVYTTDNDTQNTKFTPVGVIQGNHNEPQQWVLPNYSGRTHNVHLAPA 420

QY 421 VAPTPGQQLLFFRSTMPGCGYPNNLDCLLPQEWVQHFQEAAPASQDVALLRFVNP 480

Db 421 VAPTPGQQLLFFRSTMPGCGYPNNLDCLLPQEWVQHFQEAAPASQDVALLRFVNP 480

QY 481 TGRVLFECKLHKSGYVTVVHTGPHDLVIPPNGYFRFDSWVNOFYTLAPMNGAGRRAL 539

Db 481 TGRVLFECKLHKSGYVTVVHTGPHDLVIPPNGYFRFDSWVNOFYTLAPMNGAGRRAL 539

RESULT 14

Q917W2

ID Q917W2

AC Q917W2 PRELIMINARY; PRT; 539 AA.

DT 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Capsid protein.

GN ORF2.

OS Norwalk-like virus NLV/Burwash Landing/331/1995/US.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC Norovirus.

OX NCBI_TaxID=171857;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Hu/NLV/Burwash Landing/331/1995/US;

RX MEDLINE=97193806; PubMed=9041391;

RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;

RT "A one-tube method of reverse transcription-PCR to efficiently amplify

a 3-kilobase region from the RNA polymerase gene to the poly(A) tail

of small round-structured viruses (Norwalk-like viruses).";

RL J. Clin. Microbiol. 35:570-577(1997).


```
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Burwash Landing/331/1995/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
RA Seto Y., Monroe S.S., Glass R.I.;
RA "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995.";
RL J. Med. Virol. 53:372-383 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Burwash Landing/331/1995/US;
RX MEDLINE=9246332; PubMed=10228052;
RA Noel J.S., Fankhauser R.L., Ando T., Monroe S.S., Glass R.I.;
RA "Identification of a distinct common strain of 'Norwalk-like viruses'
RT having a global distribution.";
RL J. Infect. Dis. 179:1334-1344 (1999).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Burwash Landing/331/1995/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RA Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: AP414425; AAL13028.1; -
DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF00915; Calici_coat; 1.
SQ SEQUENCE 539 AA; 58791 MW; 57A9576B0076B5BC CRC64;

Query Match 99.0%; Score 2856; DB 12; Length 539;
Best Local Similarity 98.9%; Pred. No. 8.3e-214;
Matches 533; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGTANLVPEVNEVMALEPVVGAIAAPVAGQQNVDPWIRNNFVQAP 60
DB 1 MKMASNDANPSDGTANLVPEVNEVMALEPVVGAIAAPVAGQQNVDPWIRNNFVQAP 60

QY 61 GGEFTVSPRNAPGEILLWSAPLGLPDLNPLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
DB 61 GGEFTVSPRNAPGEILLWSAPLGLPDLNPLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120

QY 121 FAVVPPNFTPEGLSPSQVTMFPHIIIVDVRLQLEPVLPLPDRVNNFYHYNQSDNDSTIKLIA 180
DB 121 FAVVPPNFTPEGLSPSQVTMFPHIIIVDVRLQLEPVLPLPDRVNNFYHYNQSDNDSTIKLIA 180

QY 181 MLYTFLRANNAGDDVFTVSCRVLTRPSPDFDFLVPPTVESRTKPTVPILTVEEMSNS 240
DB 181 MLYTFLRANNAGDDVFTVSCRVLTRPSPDFDFLVPPTVESRTKPTVPILTVEEMSNS 240

QY 241 RFPIPLEKLYTGPSAFTVQPGRCCTTDGVLGTTQLSANTCTRGDVTTHAGSHDYT 300
DB 241 RFPIPLEKLYTGPSAFTVQPGRCCTTDGVLGTTQLSANTCTRGDVTTHAGSHDYT 300

QY 301 MNLASQWNSYDPTPEEIPAPLGTDFVGGKIQCMLTQTTREDGSTRHAKATVSTGSHVFT 360
DB 301 MNLASQWNSYDPTPEEIPAPLGTDFVGGKIQCMLTQTTREDGSTRHAKATVSTGSHVFT 360

QY 361 KLGSVQVYTTDNTNDPQTGNTKFTPVGVIGQNNHNEPQGVLPVNTSGRTGHNVLAPA 420
DB 361 KLGSVQVYTTDNTNDPQTGNTKFTPVGVIGQNNHNEPQGVLPVNTSGRTGHNVLAPA 420

QY 421 VAPTFPGEQLLFFRSTWPGCSGYENMNLDCLLPQEWQHFCOEAPAAQSDVALLRFVNP 480
DB 421 VAPTFPGEQLLFFRSTWPGCSGYENMNLDCLLPQEWQHFCOEAPAAQSDVALLRFVNP 480

QY 481 TGRVLFECKLHKSQVTVVAHTGPHDLVTPNGYFRFDSWVNWQFYTLAPMGNGAGRRAL 539
DB 481 TGRVLFECKLHKSQVTVVAHTGPHDLVTPNGYFRFDSWVNWQFYTLAPMGNGAGRRAL 539

RESULT 15
Q917W5
```

```
ID Q917W5 PRELIMINARY; PRT; 539 AA.
AC Q917W5;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Miami Beach/326/1995/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Miami Beach/326/1995/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RA "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577 (1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Miami Beach/326/1995/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
RA Seto Y., Monroe S.S., Glass R.I.;
RA "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995.";
RL J. Med. Virol. 53:372-383 (1997).
RN [3]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Miami Beach/326/1995/US;
RX MEDLINE=9246332; PubMed=10228052;
RA Noel J.S., Fankhauser R.L., Ando T., Monroe S.S., Glass R.I.;
RA "Identification of a distinct common strain of 'Norwalk-like viruses'
RT having a global distribution.";
RL J. Infect. Dis. 179:1334-1344 (1999).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Miami Beach/326/1995/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RA "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348 (2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Miami Beach/326/1995/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RA Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: AP414424; AAL13025.1; -
DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF00915; Calici_coat; 1.
SQ SEQUENCE 539 AA; 58759 MW; D6A89AF9269DCRDD CRC64;

Query Match 99.0%; Score 2856; DB 12; Length 539;
Best Local Similarity 99.1%; Pred. No. 8.3e-214;
Matches 534; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGTANLVPEVNEVMALEPVVGAIAAPVAGQQNVDPWIRNNFVQAP 60
DB 1 MKMASNDANPSDGTANLVPEVNEVMALEPVVGAIAAPVAGQQNVDPWIRNNFVQAP 60

QY 61 GGEFTVSPRNAPGEILLWSAPLGLPDLNPLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
DB 61 GGEFTVSPRNAPGEILLWSAPLGLPDLNPLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120

QY 121 FAVVPPNFTPEGLSPSQVTMFPHIIIVDVRLQLEPVLPLPDRVNNFYHYNQSDNDSTIKLIA 180
DB 121 FAVVPPNFTPEGLSPSQVTMFPHIIIVDVRLQLEPVLPLPDRVNNFYHYNQSDNDSTIKLIA 180
```


Qy	181	MLYTPLRANNAGDDVFTVSCRVLTRPSPDPDFIFLVPPTVESRTKPF	240
Db	181	MLYTPLRANNAGDDVFTVSCRVLTRPSPDPDFIFLVPPTVESRTKPF	240
Qy	241	REPIPLEKLYTGPSSAFVVPQNGRCTTDGVLGGTTQLSAVNICTFRG	300
Db	241	REPIPLEKLYTGPSSAFVVPQNGRCTTDGVLGGTTQLSAVNICTFRG	300
Qy	301	MNLASQNSNYDPTTEEIPAPLGTDPFVGKIQGMLTQTTREDGSTR	360
Db	301	MNLASQNSNYDPTTEEIPAPLGTDPFVGKIQGMLTQTTREDGSTR	360
Qy	361	KLGSVQYTTDTNDFQGTQNTKFTPVGVIQDGNHNEPQQWVLPNY	420
Db	361	KLGSVQYTTDTNDFQGTQNTKFTPVGVIQDGNHNEPQQWVLPNY	420
Qy	421	VAPTFEGEQLLFRSTMPGCSGYPNMNLDCLLPQEWVQHFCOEAAP	480
Db	421	VAPTFEGEQLLFRSTMPGCSGYPNMNLDCLLPQEWVQHFCOEAAP	480
Qy	481	TGRVLFECKLHKSGYVTVHAHTGPHDLVIPPNGYFRFDSWVNOFY	539
Db	481	TGRVLFECKLHKSGYVTVHAHTGPHDLVIPPNGYFRFDSWVNOFY	539

Search completed: June 1, 2004, 13:53:16
Job time : 34.3814 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:34:56 ; Search time 6.51141 Seconds
(without alignments)
4310.252 Million cell updates/sec

Title: US-09-926-799-5
Perfect score: 2886
Sequence: 1 MKMASNDANPSDGSSTANLVP.....VNQFTYLPAMGNGAGRRRAL 539

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2801	97.1	539	1 COAT LORDV	P54635 lordsdale v
2	1110	38.5	546	1 COAT SOUV3	Q04542 southampton
3	326	11.3	2344	1 POLN RHV	P27410 rabbit hemo
4	300.5	10.4	671	1 COAT FCVF9	P27406 feline cali
5	296.5	10.3	702	1 COAT SMSV1	P36284 san miguel
6	294	10.2	668	1 COAT FCVF4	P27405 feline cali
7	293	10.2	668	1 COAT FCVC6	P27404 feline cali
8	288.5	10.0	703	1 COAT SMSV4	P36285 san miguel
9	285	9.9	2208	1 POLN MANCV	Q69014 manchester
10	139.5	4.8	2193	1 POLG CX16T	Q9qf31 c genome po
11	133.5	4.6	2194	1 POLG HE701	P32537 h genome po
12	130.5	4.5	1007	1 Y741 CHLMU	Q9p1t6 chlamydia m
13	127.5	4.4	2193	1 POLG CX16G	Q65900 c genome po
14	127	4.4	2332	1 POLG FMDV0	P03305 f genome po
15	126.5	4.4	2333	1 POLG FMDV1	P03306 f genome po
16	122.5	4.2	2201	1 POLG CXA9	P21404 c genome po
17	122	4.2	1011	1 POLG FMDVT	P15072 foot-and-mo
18	118.5	4.1	2196	1 POLG EC05N	Q9v1j1 e genome po
19	118.5	4.1	2332	1 POLG FMDVA	P03308 f genome po
20	115	4.0	2336	1 POLG FMDVZ	P49303 f genome po
21	112.5	3.9	2175	1 POLG BOVEV	P12915 b genome po
22	112	3.9	2191	1 POLG EC06C	Q66474 e genome po
23	110	3.8	651	1 CORO YEAST	Q06440 saccharomyc
24	110	3.8	796	1 ABAA EMENI	P20945 emericeila
25	109	3.8	782	1 SMAB MOUSE	Q62179 mus musculu
26	107	3.7	2185	1 POLG CXB3W	Q66282 c genome po
27	106	3.7	1119	1 ALG3 CANAL	O74623 candida alb
28	106	3.7	2185	1 POLG CXB3N	P03313 c genome po
29	104.5	3.6	809	1 QUITA ACICA	Q59086 acinetobact
30	104.5	3.6	3487	1 CSM2 HUMAN	O72408 homo sapien
31	104	3.6	1080	1 NRK1 YEAST	P38692 saccharomyc
32	104	3.6	1196	1 AMYB PAEPO	P21543 paenibacill
33	104	3.6	1200	1 HYAL STRPU	O76536 strongyloce

RESULT 1

ID	COAT LORDV	STANDARD;	PRT;	539 AA.
AC	P54635;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	Coat protein (Capaid protein).			
OS	Lordsdale virus (Human enteric calicivirus).			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;			
OC	Norovirus.			
OX	NCBI_TaxID=82658;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96005060; PubMed=7561776;			
RA	Dingle K.E., Lambden P.R., Catl E.O., Clarke R.N.,			
RT	"Human enteric Caliciviridae: the complete genome sequence and			
RT	expression of virus-like particles from a genetic group of small			
RT	round structured viruses			
RL	J. Gen. Virol. 76:629-635 (1995).			
CC	1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.			

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch)

EMBL; X86557; CAA60255.1; InterPro; IPR004005; Calici_coat. InterPro; IPR008975; Viral_cap_coat. Pfam; PF00915; Calici_coat; 1. Coat protein; Glycoprotein. FT CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD 531 531 N-LINKED (GLCNAC. .) (POTENTIAL). SQ SEQUENCE 539 AA; 58775 MW; 585C63E7F2C5PD21 CRC64;

Query Match	97.1%;	Score	2801;	DB 1;	Length	539;			
Best Local Similarity	96.3%;	Pred. No.	1.7e-199;						
Matches	519;	Conservative	12;	Mismatches	8;	Indels	0;	Gaps	0;
QY	1	MKMASNDANPSDGSSTANLVP	EVNNEVMALEPVVGAIAAPVAGQNVQAP	60					
Db	1	MKMASNDANPSDGSSTANLVP	EVNNEVMALEPVVGAIAAPVAGQNVQAP	60					
QY	61	GGEFTVSPRNAPGEILWSAPL	PDNPYLSHLRMYNGAGFEVQVILAGNAFTAGKI	120					
Db	61	GGEFTVSPRNAPGEILWSAPL	PDNPYLSHLRMYNGAGFEVQVILAGNAFTAGKI	120					
QY	121	FAAVPPNPTEGLSPQVTMPF	HHIIVDVRQLSPVLIPLDVRRNFHHYNQSDNSTKLIA	180					
Db	121	FAAVPPNPTEGLSPQVTMPF	HHIIVDVRQLSPVLIPLDVRRNFHHYNQSDNSTKLIA	180					

QY	181	MLYTPLRANNAGDDVFTVSCRVLTRPSDPDFILVLPPTVESRTKPTVPILTVEMSNS	240
Db	181		
QY	241	RPPIPLEKLYTGPSAFVQPNQGRCTTGDVLLGTQLTSANNICTFRGDVTHIAGSHDYT	300
Db	241		
QY	301	MNLASQNSNYDPTTEIPAPLCTGPDVKGKIQGLTQTTRDGSSTRAHKATVSTGSHVHTP	360
Db	301		
QY	361	KLGSVOYTTDNNDFOTGQNTKFTPVGVIQDGNHQNPEQQWVLPNYSGRTGHNHVLAPA	420
Db	361		
QY	421	VAPTFPGEQLLFFRSTMPGSGYPNNMLDCLLPQEWVLFYQEAAPQASDVALLRFVNP	480
Db	421		
QY	481	TGRVLFCKLHKSQYTVTAHTGPHDLVIPPNGYFRFDSWVNOFYTLAPMNGAGRRRAL	539
Db	481		
RESULT 2			
ID	COAT_SOUV3	STANDARD;	PRT; 546 AA.
AC	Q04542;		
DT	01-OCT-1994 (Rel. 30, Created)		
DT	01-OCT-1994 (Rel. 30, Last sequence update)		
DT	01-OCT-1994 (Rel. 30, Last annotation update)		
DE	Coat protein (Capsid protein).		
OS	Southampton virus (serotype 3).		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;		
OC	Norovirus.		
OX	NCBI_TaxID=37129;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
EX	MEDLINE=91142023; PubMed=8380940;		
RA	Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;		
RT	"Sequence and genome organization of a human small round-structured		
RT	(Norwalk-like) virus.";		
RL	Science 259:516-519(1993).		
CC	-1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; L07418; AAA92984.1; -.		
DR	PIR; B37491; B37491.		
DR	InterPro; IPR004005; Calici_coat.		
DR	InterPro; IPR008975; Viral_Cap_coat.		
DR	Pfam; PF00915; Calici_coat; 1.		
KW	Coat protein; Glycoprotein.		
FT	CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).		
SQ	SEQUENCE 546 AA; 58774 MW; C82B2A85AD4B05EA CRC64;		
Query Match 38.5%; Score 1110; DB 1; Length 546;			
Best Local Similarity 44.6%; Pred No. 1.6e-74;			
Matches 249; Conservative 80; Mismatches 193; Indels 36; Gaps 17;			
QY	1	MKMASNDANFSDGSGT--ANLVPEVNN-EVMALEPVVGAIAAPVAGQNVLDPIRNVF	56
		: : : : : : : : : :	

CC EMBL; M67473; AAA47285.1; --
DR PIR; A41039; RWRMRH.
DR PDB; 1KHV; 16-JAN-02.
DR PDB; 1KHW; 16-JAN-02.
DR MEROPS; C24.001; --
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR004005; Calci_coat.
DR InterPro; IPR004004; Calci_pol_hel.
DR InterPro; IPR000317; Peptidase_C24.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR Pfam; PF03510; Peptidase_C24; 1.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PRO0916; 2CENDOPTASE.
DR PRINTS; PRO0918; CALICIVIRUSNS.
DR SMART; SM00382; AAA; 1.
KW Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;
KW Thiol protease; Helicase; ATP-binding; Coat protein; 3D-structure.
FT CHAIN 7506 7707 HELICASE (P2C-LIKE).
FT CHAIN 1109 1251 PROTEASE P3C.
FT CHAIN 71497 71625 RNA-DIRECTED RNA POLYMERASE.
FT NP BIND 522 529 ATP (POTENTIAL).
FT ACT_SITE 1212 1212 PROTEASE (POTENTIAL).
FT ACT_SITE 1227 1227 PROTEASE (POTENTIAL).
SQ SEQUENCE 2344 AA; 257066 MW; 1454C248F81E9212 CRC64;

Query Match 11.3%; Score 326; DB 1; Length 2344;
Best Local Similarity 25.3%; Pred. No. 1e-15;
Matches 137; Conservative 66; Mismatches 209; Indels 130; Gaps 25;

QY 4 ASNDANPSDGTANILVPEVNNVMALEVPVGAIAAIVAGQONVID---PWIRNNFVQAP 60
DB 1785 ASVPGTTTDMGPGVV--ATTSVTAENSSASIAATAGIGPPQVDQETW-RTNFYY-- 1839

QY 61 GEEFTVSPNARGEILMSAPGLDNLNLYSLHARMYNGVAGFEVQVILAGNAFTAGKII 120
DB 1840 NDVFTWADAPGSLIYTVQHSFQNNPFTAVLSQMYAGWAGMGQFRFVAGSGVGGRLV 1899

QY 121 FAAPVPNPFTEGLSPS-QVTMPHIIIVDVRQLEPVLIPLPDVRRNFYHYNQSDSTIKLI 179
DB 1900 RAVIIPGIE---IGGLEVRQFPHVIDARSLEPVITWPDRLRPNMYHTGDPGLVPTLV 1956

QY 180 AMLYPLRANNAGDDVFTVSCRVLTRPSPDPDFIFLVP--TVESRTKP--FTVPILTV 234
DB 1957 LSVYNNL-INPFGGSTSAIQVTETRPSEDFEVMIRAPSSKTVDISPAGLTTPVLTL- 2014

QY 235 EEMNSRPIPLEKLYTGSSAFVQPNQ---RCTTDGLVLTGTTQ-LSAVNICITFRGDV 290
DB 2015 -----GVGNDRWNG--QIVGLQPVPGFSTCNRHNLNGSTYGWSSPRF-----GDI 2060

QY 291 THIAGSHDYTMNLAS---QNW-----SNVD-PTEIIPALGTPD----- 325
DB 2061 DHRGASYSGSNATNLVQFVYANAGSAIDNPIQV-APDGPDMFSVPFNGPGGIPAAGW 2119

QY 326 -----FVGKIQMLTQTTRDGSTRAHKA--TVSTGS--- 355
DB 2120 VGFGALWNSGAPNVTTVQAVELGATGAPGNLQPTINTSQAQTVAKSIYAVVTGTAQN 2179

QY 356 -----VHFTPKLGSVQTTTNDNFQ-----GQNTKFTPVGVQ----- 390
DB 2180 PAGLFVMSAGIISTENASAITTPQDRIVTFTTGTPTAAAPVAKNTPIMFASVYRRTGDVN 2239

QY 391 ----DGNHQNBPQQWLPNYSRGTGNVHLAPAVAPTFGQLLFFRSTMPGCSGYPM 446
DB 2240 ATAGSANGTYGTGSOPLVTTIGLSLNNYSSA-----LMPGQ-----FFVWQLTTFASGFMEI 2291

QY 447 NL 448
DB 2292 GL 2293

RESULT 4
COAT_FCVF9 STANDARD; PRT; 671 AA.
AC P27406;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Coat protein (Capsid protein).
GN Cpp76.
OS Feline calicivirus (strain F9) (FCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=11981;
RN [1]_TaxID=11981;
RP SEQUENCE FROM N.A.
RX MEDLINE=92410623; PubMed=1529544;
RA Carter M.J., Milton I.D., Meanger J., Bennett M., Gaskell R.M.,
RA Turner P.C.;
FT "The complete nucleotide sequence of a feline calicivirus.";
RL Virology 190:443-448(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92339673; PubMed=1633955;
RA Meanger J., Carter M.J., Gaskell R.M., Turner P.C.;
FT "Cloning and sequence determination of the feline calicivirus strain
FT F9.";
RL Biochem. Soc. Trans. 20:265-26S(1992).
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M86379; AAA79327.1; --
DR EMBL; Z11536; CAA77636.1; --
DR PIR; B43382; VCMWF9.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 671 AA; 73441 MW; 33BE86D8370D5ES CRC64;

Query Match 10.4%; Score 300.5; DB 1; Length 671;
Best Local Similarity 25.3%; Pred. No. 1.4e-14;
Matches 141; Conservative 80; Mismatches 238; Indels 99; Gaps 28;

QY 12 DGSITANLVPEVNNV--MALEPVVGAIAAIVAGQONVIDPWIRNNFVQAPGGETV--- 66
DB 127 DGSIT--APEQGTVMVGVIAPSAQMSTAADATKGSVDSEW-----EAFPSHTSVNW 178

QY 67 SPRNAPGEILMSAPGLDNLNLYSLHARMYNGVAGFEVQVILAGNAFTAGKIIFAAVPP 126
DB 179 STSETQKILPKQSLGSLNPLNLYSLHAKLYVAVWSGSIERSISGSGVFGKLAIVVPP 238

QY 127 NPTTEGLSPSQVT---MEPHIIVDVRQLEPVLIPLPDVRRNFYHYNQSDSTIKLIAMLY 183
DB 239 -----GVDPVQSTMLQYPHVLFDAQVEPVIFCLPDLRSTLYHLMSDIT-SLVINVY 292

QY 184 TPLRANNAGDDVFTVSC--RVLTRPSDDFIFLVPPTVESRTKPTFTVPILTVEMSNSR 241

Db 293 NDL-INPYANDANSSGCIYVETKPGDPKFKHLLKPP--GSMHTGSIIPSDLIPTKTS-- 347
QY 242 FPIPLEKLTGSPSSAFVWOP--QNGR-----CTTDGVLGGTQLSAVNICTFRGDVTHI 293
Db 348 --LWGNRWSDITDFVIRPFVQANRHFDFNQETAGNTPFRFPISVITTEQNGAKLGI 405
QY 294 AGSHDYTMNLASQNSNYD-PTEETAPLGTDFVQKIQGMLTQTTREDGSTR--HKAT 350
Db 406 GVATDVIYVPCIPDGMDDTTPGELIPA---GDYA-----ITNGTGNDITATGYDTAD 455
QY 351 VSTGSHVF-----TPKLGSVQYTT-----DTNDDFTGQNTKFTPVGVQI 390
Db 456 IKNNTNFRGMYICGSLORAGDKKISNTAFITATLDGNNKKNPCNTIDQSKIIVFQ 515
QY 391 D--GNNHQNQEPQWVLPNYSG---RTGHNHVLAPAVAPTFF-----GEQLLFFRSTM 437
Db 516 DNHVGKAKTSDTLALLGYTGIGEAIGSDRDRVVRIS-TLPETGARGNHPHFVNSI 574
QY 438 PGCSGYNNMLDCLLPQWVHFCQAAA-----PAQSDVALLRFVNPDTGRVLPCKLH 491
Db 575 K--LGVVIRSIDVFNISQ--ILHTSRQLSLNHYLLPPDS-FAVYRII--DSNGSWDFDIGID 627
QY 492 KSGVTVTAHTGPHDLVIP 509
Db 628 SDGFSFVGVSGFGKLEFP 645

RESULT 5
COAT_SMSV1 STANDARD; PRT; 702 AA.
AC P36284;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Coat protein (Capsid protein).
OS San Miguel sea lion virus (serotype 1) (SMSV 1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesiviruses.
OX NCBI_TaxID=36406;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92410750; PubMed=1529644;
RA Neill J.D.;
RT "Nucleotide sequence of the capsid protein gene of two serotypes of
RT San Miguel sea lion virus: identification of conserved and non-
RT conserved amino acid sequences among calicivirus capsid proteins.";
RL Virus Res. 24:211-222(1992).
CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M87481; AAA16217.1; -
DR PIR; A48562; A48562.
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 493 493 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 702 AA; 77850 MW; E6E5A58523DEE3D7 CRC64;

Query Match 10.3%; Score 296.5; DB 1; Length 702;
Best Local Similarity 26.1%; Pred. No. 2.9e-14;
Matches 141; Conservative 72; Mismatches 199; Indels 129; Gaps 30;

QY 11 SDG-STANLYPEVNEVMALPVVG---AAIAAPVAGQONVIDPWIRNFWOAPGGEFT 65
Db 153 SDGPGGADIYTERGCTVVOQOPVPAQASALTLLAASTGK--TVD-----CEWT 198
QY 66 V-----SPRNAGEILMSAPLGPDLNPNLYSHLARMYNGYAGGFVQVVLGNAFT 115
Db 199 TFFSYHTAVNWSYTEAQKILFSLRSLSPNLPYLRHISLYSSGSGIDVRFVTSQGVF 258
QY 116 AKLIIFAAPVDPNF-PTGELSPSQVMTWPHHIIIVDROLEPVLILPLDPVRNNFYHNSQNS 174
Db 259 GGLAALIVPPIEGIEFVE--SPTML-QYPHVLFPARQTEPVIFIPDIRKTLYHSMDDTDT 315
QY 175 TIKLIAMLYTPL-RANNAGDDVFTVSCRVLTRFSPDFEFILVPPTVSESTKFTFPILT 233
Db 316 T-RLVIVYVNELINPYEQSEPKSCSITVETRPSSDFTSLKKPP--GSLKHGSIPSDL 372
QY 234 VEEMSNRFPPIPLEKLYTGPSSAFVPOQ-----NGRCTTGVLLGTQLSAVNICTFRGD 289
Db 373 IPR--NSRH--WMGNRWMTIDGFVQPRVQFQSNRHFDFDSTTTGWTSTPYIPI----- 422
QY 290 VTHIAGSHDYTMNLASQNSNYDPTB---IPA-PLGTDPDFVGIQGMILTQTTREDGS- 343
Db 423 -----EVTLEKLDGGQYFKVTDTEKSLVPLGPDGWDTT-----IPTAMTASNGNY 469
QY 344 --TRAHKATVSTGSHVF-----TPKLGSVQYTTDTNNDFOTGONTKFTPVGVQDG 392
Db 470 DYTVAEYRI-TNNGT-HFKGFYIMGNLTTKVKG-----SDNLGETQTSRTILFASVGNKQD 524
QY 393 N-----NHQNEPQWVLPNYSGRTHNHLAPAVAPTFFGEQLLFFRSTMPGSCGYNNML 448
Db 525 NTINPTKITSLSLVVD-----ANNVSAATAKTTT-----WHSTM-----SHL 563
QY 449 DCLLPQWVHFCQEAAPQSD-----VALL--RFVNPDTGRVLFCKLHSGYVTVAH 500
Db 564 GVLVDE-----SPVGSDSIKVYRIATLPEAFTNGGPNFPFTNKI-QIGHFRAH 613
QY 501 T 501
Db 614 T 614

RESULT 6
COAT_FCVF4 STANDARD; PRT; 668 AA.
AC P27405;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Coat protein (Capsid protein).
OS Feline calicivirus (strain Japanese F4) (FCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesiviruses.
OX NCBI_TaxID=11980;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91306470; PubMed=1853578;
RA Tohya Y., Taniguchi Y., Takahashi E., Utagawa E., Takeda N.,
RA Miyamura K., Yamazaki S., Mikami T.;
RT "Sequence analysis of the 3'-end of feline calicivirus genome.";
RL Virology 183:810-814(1991).
CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90357; BAA14371.1; -
DR PIR; B40481; VCMWFC.

```

DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 668 AA; 73569 MW; 85BDCB85804E503 CRC64;

Query Match 10.2%; Score 294; DB 1; Length 668;
Best Local Similarity 24.6%; Pred. No. 4.2e-14;
Matches 120; Conservative 61; Mismatches 171; Indels 136; Gaps 21;

QY 12 DGSTANLVPENVNEV--MALEPVVGAATAAPVAGQNVDPWIRNPNVQAPGGEFTV--- 66
DB 127 DGSIT--TPEQGTAVGVIAEPAQMSAQMSTADWASKSVDSW-----EAFSFTSVNW 178

QY 67 SPRNAPGEILMSAPLGPDLNPNYLSHARMYNGYAGGFEVQVILAGNAFTAGKIIFAAVPP 126
DB 179 STSETQKILFKQSLGPLNPNYLHSLKLYVAMSGSIEVRFSISGSGVFGKLAIVVPP 238

QY 127 NPTTEGLSPSQVT---MPPHIIIVDVRQLEPVLPLPDVRRNPNFYHYNQNSDKILAMLY 183
DB 239 -----GVDPVQSTSLMQLYPHVLFDARQVEPFIPTDLRSTLYHVMSTDTT-SLVIMVY 292

QY 184 TPLRANNAGDDVFTVSC--RVLTSPSPDFEFLVPP-----TVES----- 222
DB 293 NDL--INPYANDSNSSGCIIVTETKPGDFKPHLLKPPGSLVTHGSIPLSDLPKSSSLWIG 351

QY 223 -----RTKPFVPIILTVSEMSNRFPPIPLE 247
DB 352 NRYWTDITDFVIRPFVQANRHFDFNQETAGSTFRPRITI--TISEKNSKLGIGVA 408

QY 248 KLYTGP-----SSAFVQPNQGRCTTGDVLLGTQ---LSAV--NICTPRGDTV 291
DB 409 TDYIIPGIPDGPWDTTADKLIPAGDYSITTTGEGNDIKTAQAYDTAAVYVNTTFRG--M 466

QY 292 HIAGSHDYTMNLASQNSNYDPTTEIPAPLGTGPDVFGKIQGLMTOTTR-EDGSTRAHKAT 350
DB 467 YICGS-----LQAWGD-----KKISNTAFITTAIRDGNEIKPSNT 502

QY 351 VSTGSHVFTPKLGSQYITTTNNDFQTGONT-----KFTPVGVIOGNNHQNPEQWVLPN 406
DB 503 IDMTKL-----AVYQDTHVEQEVQTSDDTLALLGYTGIGEEAIGSNDRVVRISVLPE 555

QY 407 YSGRTGHN 414
DB 556 AGARGGNH 563

RESULT 7
COAT_FCV66 STANDARD; PRT; 668 AA.
ID COAT_FCV66 AC P27404;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein (Capsid protein).
OS Feline calicivirus (strain CFI/68 FIV) (FCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=11979;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=91374597; PubMed=1716692;
RA Neill J.D., Reardon I.M., Heinrikson R.L.;
RT "Nucleotide sequence and expression of the capsid protein gene of
RT feline calicivirus.";
RL J. Virol. 65:5440-5447(1991).
RN [2]

```

```

RP SEQUENCE FROM N.A.
RA Neill J.D.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M32819; AAA42925.1; -.
CC EMBL; U13992; AAC13993.1; -.
DR PIR; A40507; VCMWFF.
DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 668 AA; 73550 MW; 9E52312108D4D41 CRC64;

Query Match 10.2%; Score 293; DB 1; Length 668;
Best Local Similarity 25.3%; Pred. No. 5e-14;
Matches 117; Conservative 66; Mismatches 195; Indels 84; Gaps 18;

QY 12 DGSTANLVPENVNEV--MALEPVVGAATAAPVAGQNVDPWIRNPNVQAPGGEFTV--- 66
DB 127 DGSIT--TPEQGTAVGVIAEPAQMSAQMSTADWASKSVDSW-----EAFSFTSVNW 178

QY 67 SPRNAPGEILMSAPLGPDLNPNYLSHARMYNGYAGGFEVQVILAGNAFTAGKIIFAAVPP 126
DB 179 STSETQKILFKQSLGPLNPNYLHSLKLYVAMSGSVDVRFSGSGVFGKLAIVVPP 238

QY 127 NPTTEGLSPSQVT---MPPHIIIVDVRQLEPVLPLPDVRRNPNFYHYNQNSDKILAMLY 183
DB 239 -----GIDPVQSTSLMQLYPHVLFDARQVEPFIPTDLRSTLYHVMSTDTT-SLVIMVY 292

QY 184 TPLRANNAGDDVFTVSC--RVLTSPSPDFEFLVPPVEGRKDFVPIILTVSEMSNR 241
DB 293 NDL--INPYANDSNSSGCIIVTETKPGDFKPHLLKPP--GSMITHGSIPLSDLPKSSS--- 347

QY 242 FPIPLEKLYTGPFSSAFVQVP---QNGR-----CTTDGVLGTTQLSAVNICTFRGDVTHI 293
DB 348 --LMIGNRFWSDIITDFVIRPFVQANRHFDFNQETAGSTFRPRITITISVKESAKLGI 405

QY 294 AGSHDYTMNLASQNSNYDPTTEIP---APLGTDPFVGKIQGLMTQTTREDGSTR----- 345
DB 406 GVATDYIVPGIPDGN---PDTTIPGELVPVGDYAITNGTNNDDITTAQAYDAATEIRNNT 461

QY 346 -----AHKATVSTGSHVFTPKLGS-----VQYTTDNNDFQ 376
DB 462 NFRGMYICGSLQRAWGDKKISNTAFITTTGVDGAKLIPSNITIDQKIAVQDTHANKHVQ 521

QY 377 TQONT----KFTPVGVIOGNNHQNPEQWVLPNYSGRGTGHN 414
DB 522 TSDDTLALLGYTGIGEEAIGADRVRVIRISVLPERGARGGNH 563

.. RESULT 8
COAT_SMSV4
ID COAT_SMSV4 STANDARD; PRT; 703 AA.
AC P36285;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)

```

Coat protein (Capsid protein).
 S San Miguel sea lion virus (serotype 4) (SMSV 4).
 C Viruses; sRNA positive-strand viruses, no DNA stage; Caliciviridae;
 C Vesivirus.
 X NCBI_TaxID=36407;
 [1]
 SEQUENCE FROM N.A.
 MEDLINE=92410750; PubMed=1529644;
 A Neill J.D.;
 T "Nucleotide sequence of the capsid protein gene of two serotypes of
 T San Miguel sea lion virus: identification of conserved and non-
 T conserved amino acid sequences among calicivirus capsid proteins."
 L Virus Res. 24:211-222(1992).
 C -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 C -----
 C This SWISS-PROT entry is copyright. It is produced through a collaboration
 C between the Swiss Institute of Bioinformatics and the EMBL outstation -
 C the European Bioinformatics Institute. There are no restrictions on its
 C use by non-profit institutions as long as its content is in no way
 C modified and this statement is not removed. Usage by and for commercial
 C entities requires a license agreement (See http://www.isb-sib.ch/announce/
 C or send an email to license@isb-sib.ch).
 C -----
 C EMBL; M87482; AAA16220.1; -;
 C PIR; C48562; C48562; Calici_coat.
 C InterPro; IPR008975; Viral_cap_coat.
 C Pfam; PF00915; Calici_coat; 1.
 C Coat protein; Glycoprotein.
 C CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).
 C CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
 C CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
 C CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).
 C CARBOHYD 482 482 N-LINKED (GLCNAC. .) (POTENTIAL).
 C SEQUENCE 703 AA; 77721 MW; C5DAD8223B261073 CRC64;
 C -----
 C Query Match 10.0%; Score 288.5; DB 1; Length 703;
 C Best Local Similarity 27.3%; Pred. No. 1.1e-13;
 C Matches 101; Conservative 56; Mismatches 164; Indels 49; Gaps 16;
 C
 C 11 SDG-STANLVPEVNNEWALEPV---VGAIAAPVAGQQNVDPWIRNNFVQAPGGEFTV 66
 C 153 SDGSGSAEIVTEQGTVVQQQPAPAPATLATALTAGSKSVEQEMWTFYSYTSINWSTV 212
 C 67 SPRNAPCEILWSAPGLDLPNLYSLHARMYNGVAGFEVQVILAGNAFTAGKIIIFAAVPP 126
 C 213 ---ESQKILYSQALNSINPYLDHIAKLYSTWSGGIDRVFTVSGVFGGKLAALLVPP 269
 C 127 NF-PTGELSPSQVTFPHIIVDVRLQPLVPLIPDVRRNFYHYNQSDSTIKLIAMLYTP 185
 C 270 GVEPIESVSMQ---YPHVLFDARQTEFEVFTIPDTRKTLFHSMDTDTT-KLVINPY-- 323
 C 186 LRANAGDDVFTVSCRVLTRPSPDFDFIFVPP--TVESRTKPF-TVPILTVEEMSNRF 242
 C 324 ---ENGVENKTCITVETRESADFTALLKPPGLSLIKHGSIPSDLIPRNSAHWMGNR-- 378
 C 243 PIPLEKLYTGPSAFVQVQ---NGRCTTDGVLGTTQLSANNI-CTFGDVTTHAGSH 297
 C 379 -----WWSITSGFSVQPRVFSQNRHFDFTSTTTGWSSTPYYPVPIEIKIGKV----GSN 427
 C 238 DYTWNLASQNSNYDPTEEIPALGTDPFGVKIQGMLTQTTREDGSGTRAHKATVSTGSHV 357
 C 428 N-----KNFHVLDTDKALVP-GIPD--GWPTTTPIDETKATNGNFSYGSVRAGSTT 476
 C 358 FTPLKLGVSQY 367
 C 477 IKPENSTHF 486
 C
 C RESULT 9
 C POLN MANCV STANDARD; PRT; 2208 AA.
 C ID ID
 C AC Q69014;

Db 450 ATGKMLIAYTPPGNVPADRIITAMLTG---HVWDFGLQSSVTLVVPWISNTHY----- 500
QY 173 DSTIKLIAMLYPLRAN-NAGDDVFTSVCLTRSPDFFDFILVPPVTESRTKPFVPI 231
Db 501 -----RAHARAG-----YFDY-----YTTGI 516
QY 232 LTVEMSNRRPPIPLEKLYTGSSAFVQVQPNQRCRTTGGVLLGTQLSNAVICTFRGDTV 291
Db 517 IITVQTNVVPFI-----GAPTAYIVA----- 539
QY 292 HIAGSHD-YTNKLASQNSNYDPTTEIPAPLGTDPFVGKIQ-----MLTOTTRED----- 341
Db 540 -LAAAOQDNFTMKLXD-----TEDI-----EGTANIQDPIADMDIQVNNQVNR 584
QY 342 -----GSTRAHKATVSGVHFTPKLGSVQVTTDTNDFOTGQNTKFTFPVGVQID 391
Db 585 LTAQLVPLTAADTEASSHRLGTGW--PALQAA-----ETGASSNASDKNLIET 631
QY 392 G---NNHQEPQOWLPNVSGRTGHNVHLAPAVTFFGEOQLLFRSTMPGCGYPNNML 448
Db 632 RCVLNHST--OETAIGNFFSRAG-----LVSITMTPTGTQ-----NTDGVNWDI 676
QY 449 DCL 451
Db 677 DLM 679

RESULT 11
POLG HE701
ID -POLG HE701 STANDARD; PRT; 2194 AA.
AC P32537;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome protein [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VP6 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE Human enterovirus 70 (strain J670/71) (EV 70).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.
OC NCBI_TaxID=31915;
OX [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=91037960; PubMed=2172447;
RA Ryan M.D., Jenkins O., Hughes P.J., Brown A., Knowles N.J., Booth D., Minor P.D., Almond J.W.;
RT "The complete nucleotide sequence of enterovirus type 70: relationships with other members of the picornaviridae.";
RL J. Gen. Virol. 71:2291-2299(1990).
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the poliovirus polypeptide. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA} (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; D00820; BAA18891.1; -.

DR PIR: A36253; GNNYET.
DR HSP: P03300; IPOV.
DR MEROPS; C03.UBA; -.
DR MEROPS; C03.UBB; -.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR004004; Calici_pol_hel.
DR InterPro: IPR009003; Cys_Ser_trypsin.
DR InterPro: IPR001199; Pept_3C_picorn.
DR InterPro: IPR000081; Peptidase_C3.
DR InterPro: IPR003138; Pico_P1A.
DR InterPro: IPR002527; Pico_P2B.
DR InterPro: IPR001676; Rhv.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF00548; Cys-protease_3C; 1.
DR Pfam: PF02226; Pico_P1A; 1.
DR Pfam: PF00947; Pico_P2A; 1.
DR Pfam: PF01552; Pico_P2B; 1.
DR Pfam: PF00073; rhv_3.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICVIRUSNS.
DR ProDom: PD001125; Cys_protease_3C; 1.
DR ProDom: PD001306; Pico_P2A; 1.
DR ProDom: PD001274; Pico_P2B; 1.
DR SMART; SM00382; AAA_1.
KW Polyprotein; Coat protein; Core protein; Transferase; RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate; Lipoprotein.
KW CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 319 COAT PROTEIN VP2.
FT CHAIN 320 561 COAT PROTEIN VP3.
FT CHAIN 562 871 COAT PROTEIN VP1.
FT CHAIN 872 1014 CORE PROTEIN P2A.
FT CHAIN 1015 1113 CORE PROTEIN P2B.
FT CHAIN 1114 1443 CORE PROTEIN P2C.
FT CHAIN 1444 1532 CORE PROTEIN P3A.
FT CHAIN 1533 1554 GENOME-LINKED PROTEIN VP6.
FT CHAIN 1555 1737 PICORNAIN 3C.
FT CHAIN 1738 2194 RNA-DIRECTED RNA POLYMERASE P3D.
FT LIPID 2 N-myristoyl glycine (in host) (By similarity).
FT ACT SITE 1701 PROTEASE (POTENTIAL).
FT ACT SITE 1715 PROTEASE (POTENTIAL).
SQ SEQUENCE 2194 AA; 244590 MW; 15DBAE96EE06673C CRC64;
Query Match 4.6%; Score 133.5; DB 1; Length 2194;
Best Local Similarity 23.7%; Pred. No. 0.16; Indels 47; Gaps 9;
Matches 57; Conservative 34; Mismatches 103; Gaps 9;
QY 54 NNFVQAPGGE---FTVSPRNAPGELIWSAPL-----GPDLPYLSHLMNNGYAGGFV 105
Db 375 NNVDASGVERLRVQISQSDMDQLLENIPLDIQLEGLPLNTLLGNISRYTHWSSLE 434
QY 106 QVILAGNAFTAGKIIPAAVPPNPTGELSPS---QVTMPFHIIIVDRQLEPVLIPDVR 162
Db 435 TFMFCGFSMTTGLIICITPPG---GSSPTDRMQAMLAATHVVDVDFGLQSSITIIIPWIS 490
QY 163 NNFHYNQNSDSTIK-----LIAMLYTPLRAN-NAGDDVFTVSCRV-----LTPSP 208
Db 491 GSHYRMFTDAKAINANVGVVTCFMQTNLVAPVGAADQCIYGVMAAKKDFNLMRD 550
QY 209 DF-----DFIFLVPTVESRTKP---FTVPILTVEEMSNRRPPIPLEKLY 250
Db 551 DIQSAILPEQAATTQIGIVKTVANTVESEIKAEIGVIFPSLNAVETGATSNTEPSEA 610
QY 251 T 251
Db 611 T 611

```
RESULT 12
Y741_CHLMU STANDARD; PRT; 1007 AA.
AC Q9PJT6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TC0741 precursor.
GN TC0741.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN0572/CT456/TC0741
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE002342; AAF39550.1; -.
CC PIR; H81670; H81670.
CC TIGR; TC0741; -.
CC SIGNAL 1 51
CC CHAIN 52 1007 HYPOTHETICAL PROTEIN TC0741.
CC SEQUENCE 1007 AA; 104006 MW; 842800C0871B1518 CRC64;
CC
CC Query Match
CC Best Local Similarity 4.5%; Score 130.5; DB 1; Length 1007;
CC Matches 75; Conservative 39; Mismatches 128; Indels 79; Gaps 14;
QY 219 TVESRTKPTVPIITVVEENSGRPIPLEKLYTGSSAFVQVQNGRCCTTDGVLGTTQL 278
DB 2 TTPISNPSPIPTVTVTSTTASSGSLGTSTVSTSTSTVAQTAT---TTSSASTSIQS 58
QY 279 SAVNICTFEGDVTIAGSHDVTWNLASQWNSYDTEETPAFLGTPDFVGKIQGMLTQT 338
DB 59 SGENIQSTGTSPIT-----SSVSTAPSASATANKTSVAVGSKI---TSOBT 106
QY 339 RDGSGTRAHKAIVTSGSVHF-----TPKLGs-----VQYTTDNDNFOTGN 380
DB 107 SESET---QATSDGEVSSNDVDVTPNNSDSTVSDSYQDVETQYKTIENN---GEN 159
QY 381 TKFTPVGVIQDGNHQNHPQWLVNYSRGTGNVH-----LAPAVAPTFPGEQLLF 432
DB 160 T-YETIGSGEKNTVQVE-----SHASGTGNPINNQEAIRQLRSSTYTSPPRNIIF 211
QY 433 FRSTWPGSGYNNMLDCLLPQEWVQHFQCEAPQAQSDVALLRFV-NPDTGRVLFECKLH 491
DB 212 ----SPGPEGLNMLSPSPSPD-----KSLLAFLSNPNTKAMLE----- 249
QY 492 KSGYVTVATGPHDLVIPNG 512
DB 250 HSHLVFIDTTRSSFIYPNG 270
```

```
RESULT 13
POLG_CX16G STANDARD; PRT; 2193 AA.
AC Q65900;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (P1a); Coat protein VP2
DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein
DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-
DE linked protein VPG (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].
OS Coxsackievirus A16 (strain G-10).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=69159;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94303216; PubMed=8030260;
RA Poyry T., Hyypia T., Horsnell C., Kinnunen L., Hovi T., Stanway G.;
RT "Molecular analysis of coxsackievirus A16 reveals a new genetic group
RT of enteroviruses."
RL Virology 202:982-987(1994).
CC -!- FUNCTION: It is thought that the P2C protein attaches to vesicular
CC membranes and is associated with viral RNA synthesis.
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA] (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U05876; AAA50478.1; -.
CC HSPF; P03300; IPOV.
CC MEROPS; C03.022; -.
CC InterPro; IPR003593; AAA ATPase.
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR000399; Pept_3C_picorn.
CC InterPro; IPR000081; Peptidase_C3.
CC InterPro; IPR003138; Pico_P1A.
CC InterPro; IPR002527; Pico_P2B.
CC InterPro; IPR01676; Rhv.
CC InterPro; IPR006005; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_Psvir.
CC InterPro; IPR008975; Viral_cap_covst.
CC Pfam; PF00548; Cys-protease-3C; 1.
CC Pfam; PF02226; Pico_P1A; 1.
CC Pfam; PF00947; Pico_P2A; 1.
CC Pfam; PF01552; Pico_P2B; 1.
CC Pfam; PF00073; rhv; 3.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICVIRUSNS.
CC ProDom; PD001125; Cys_protease_3C; 1.
CC ProDom; PD001306; Pico_P2A; 1.
```

OC Aphthovirus
ON NCBI_TaxID=73482;
RX [1]
ON SEQUENCE FROM N.A.
RP STRAIN=O1K;
RC STRAIN=84297249; PubMed=6089122;
RA MEDLINE=84297249; PubMed=6089122;
RX Forsa S., Strabel K., Beck E., Schaller H.;
RA "Nucleotide sequence and genome organization of foot-and-mouth
RT disease virus.";
RL Nucleic Acids Res. 12:6587-6601(1984).
RL [2]
RN SEQUENCE FROM N.A.
RN STRAIN=O1BFS;
RX STRAIN=83143292; PubMed=6298715;
RC MEDLINE=83143292; PubMed=6298715;
RX Makoff A.J., Paynter C.A., Rowlands D.J., Boothroyd J.C.;
RA "Comparison of the amino acid sequence of the major immunogen from
RT three serotypes of foot and mouth disease virus.";
RL Nucleic Acids Res. 10:8285-8295(1982).
RL [3]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX MEDLINE=89143740; PubMed=2537470;
RA Acharya R., Fry E., Stuart D., Fox G., Rowlands D., Brown F.;
RX "The three-dimensional structure of foot-and-mouth disease virus at
RT 2.9-A resolution.";
RL Nature 337:709-716(1989).
RL CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA] (N).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- MISCELLANEOUS: THE STRAIN O1K SEQUENCE IS SHOWN.
CC -1- MISCELLANEOUS: THE COAT PROTEIN VP1 CONTAINS THE MAIN ANTIGENIC
CC DETERMINANTS OF THE VIRION; THEREFORE, CHANGES IN ITS SEQUENCE
CC MUST BE RESPONSIBLE FOR THE HIGH ANTIGENIC VARIABILITY OF THE
CC VIRUS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X00871; CAA25416.1; -;
DR EMBL; J02185; AAA42635.1; -;
DR PDB; 1QMY; 18-SEP-01.
DR PDB; 1QOL; 18-NOV-00.
DR MEROPS; C03.008; -;
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR004080; FMDVp1coat.
DR InterPro; IPR008739; Peptidase_C28.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_Psvir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF05408; Peptidase_C28; 1.
DR Pfam; PF00073; rhv; 3
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICVIRUSNS.
DR PRINTS; PR01542; FMDVp1COAT.
KW Polyprotein; Coat protein; Core protein; Core protein; RNA-directed RNA polymerase;
KW Transferase; Hydrolase; Thiol protease; Nonstructural protein;
KW Ynstate; 3D-structure; Lipoprotein.
KW CHAIN 1 201 NONSTRUCTURAL PROTEIN P20A.
FT CHAIN 202 286 COAT PROTEIN VP4.
FT CHAIN 287 504 COAT PROTEIN VP2.

FT CHAIN 505 724 COAT PROTEIN VP3.
 FT CHAIN 725 937 COAT PROTEIN VP1.
 FT CHAIN 938 1107 CORE PROTEIN P12.
 FT CHAIN 1108 1425 CORE PROTEIN P34.
 FT CHAIN 1426 1578 CORE PROTEIN P14.
 FT CHAIN 1579 1649 GENOME-LINKED PROTEIN VP3.
 FT CHAIN 1650 1862 PROTEASE.
 FT CHAIN 1863 2332 RNA-DIRECTED RNA POLYMERASE.
 FT LIPID 202 202 N-myristoyl glycine (in host).
 FT DISULFID 511 511 INTERCHAIN (IN VP3 DIMER).
 FT DISULFID 406 858 IN VP2-VP1 DIMER.
 FT VARIANT 780 780 I -> V (IN STRAIN O1BFS).
 FT VARIANT 808 808 G -> R (IN STRAIN O1BFS).
 FT VARIANT 861 861 N -> S (IN STRAIN O1BFS).
 SQ SEQUENCE 2332 AA; 258924 MW; 4A83176F43447D68 CRC64;

Query Match 4.4%; Score 127; DB 1; Length 2332;
 Best Local Similarity 22.3%; Pred. No. 0.52;
 Matches 73; Conservative 39; Mismatches 134; Indels 82; Gaps 14;

QY 92 LARMYNGYAGGEVQVILAGNAFTAGKIIFAAVPPNFTPEGLSPQVTFPHIIVDVROL 151
 DB 380 LTDSYAYMNGWDVEVTAQVNGFNGGCLLVAMVPELYSQRKELYQLTFPHQFINPRTN 439
 QY 152 EPLVPLDVRNRYFY-HYNQSDSTIKLFIAMLYTHLRANNAGDDVFTVSCRVLTRPSDF 210
 DB 440 MTAHITVPFVGVRNYDQYKHPWT--LVVMVAPLTVNTEG-----APQI 483
 QY 211 DIFVLVPPVESRTKPTFTVILTVEMSNRRPIPLEKLY-----TGPSSAFVW----- 259
 DB 484 KYANIAPT-----NVHAGEPPSKEGIFPVACSDGYGGLVTTDPKTADPVYGVKVN 535
 QY 260 QPQN---GRCTDGVLLGTTQLSANNICTFRGDVTHIAGSHD-----YTNLASQNMWS 309
 DB 536 PPRNQLPGRFTN---LLDVAE-ACPTFLFEGGVPTVTKTDSRDLAQFDMSLAAKQMS 591
 QY 310 NYDPTTEIPAPLGPDPFVKIOGMLTQTTRDGSTRAHKATVSTG-----SVHFTPK 361
 DB 592 N-----TFLAGLAQYTYQYSGTINLH--FMFTGTDAKARYMAYAPP 632
 QY 362 LGSVQVYTTDT-----NNDFQTGONTKFT 384
 DB 633 GMEPKPTPEAAAHCHIAEWDGTGLNSKFT 660

RESULT 15

POLG_FMDV1
 ID POLG_FMDV1 STANDARD; PRT; 2333 AA.
 AC P03306; Q84768; Q84750; Q84751; Q84752; Q84753; Q84754; Q84760;
 AC Q84761; Q84762; Q84763; Q84764; Q84765; Q84766; Q84767; Q84768;
 AC Q84769; Q89824;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat
 proteins VP1 TO VP4; Coat protein P52; Genome-linked proteins VP31 TO
 DE VP33; Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed
 DE RNA polymerase P56A (EC 2.7.7.48)].
 OS Foot-and-mouth disease virus (strain A10-61) (Aphthovirus A) (FMDV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Aphthovirus.
 OK NCBI_TaxID=12112;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84169547; PubMed=6324120;
 RA Carroll A.R., Rowlands D.J., Clarke B.E.;
 RT "The complete nucleotide sequence of the RNA coding for the primary
 RT translation product of foot and mouth disease virus.";
 RL Nucleic Acids Res. 12:2461-2472 (1984).
 RN [2]
 RP SEQUENCE OF 115-1048 FROM N.A.
 RX MEDLINE=82211814; PubMed=6282711;
 RA Boothroyd J.C., Harris T.J.R., Rowlands D.J., Lowe P.A.;

RT "The nucleotide sequence of cDNA coding for the structural proteins
 of foot-and-mouth disease virus.";
 RL Gene 17:153-161 (1982).
 CC -|- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the
 CC poliovirus polyprotein. In other picornavirus reactions Glu may be
 CC substituted for Gln, and Ser or Thr for Gly.
 CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA} (N).
 CC -|- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -|- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; V01130; CAA24361.1; -;
 CC EMBL; X00429; CAA25127.1; -;
 CC MEROPS; C03.008; -;
 CC MEROPS; C28.001; -;
 CC InterPro; IPR004004; Calici_pol_hel.
 CC InterPro; IPR009003; Cys_Ser_trypsin.
 CC InterPro; IPR004080; FMDVP1coat.
 CC InterPro; IPR008739; Peptidase_C28.
 CC InterPro; IPR001676; Rhv.
 CC InterPro; IPR000605; RNA_helicase.
 CC InterPro; IPR007095; RNA_pol_DS_PS.
 CC InterPro; IPR001205; RNA_pol_P3D.
 CC InterPro; IPR007094; RNA_pol_PSVir.
 CC InterPro; IPR008975; Viral_cap_coat.
 CC Pfam; PF05408; Peptidase_C28; 1.
 CC Pfam; PF00073; rhv; 3.
 CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
 CC Pfam; PF00910; RNA_helicase; 1.
 CC PRINTS; PR00918; CALICIVIRUSNS.
 CC PRINTS; PR01542; FMDVP1COAT.
 CC Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;
 CC Transferase; Hydrolase; Thiol protease; Nonstructural protein;
 CC Myristate; Lipoprotein.
 CC CHAIN 1 201 NONSTRUCTURAL PROTEIN P20A.
 CC CHAIN 202 286 COAT PROTEIN VP4.
 CC CHAIN 287 504 COAT PROTEIN VP2.
 CC CHAIN 505 725 COAT PROTEIN VP1.
 CC CHAIN 726 937 COAT PROTEIN VP3.
 CC CHAIN 938 1578 CORE PROTEIN P52.
 CC CHAIN 1579 1601 GENOME-LINKED PROTEIN VP31.
 CC CHAIN 1602 1625 GENOME-LINKED PROTEIN VP32.
 CC CHAIN 1626 1649 GENOME-LINKED PROTEIN VP33.
 CC CHAIN 1650 1863 PROTEASE P20B.
 CC CHAIN 1864 2333 RNA-DIRECTED RNA POLYMERASE P56A.
 CC CHAIN 202 202 N-myristoyl glycine (in host).
 CC CHAIN 202 396 S -> C (IN REF. 2).
 CC CHAIN 396 632 P -> L (IN REF. 2).
 CC CHAIN 632 632
 CC CHAIN 632 632
 CC CHAIN 2333 AA; 259645 MW; 4FC667DCC521BC60 CRC64;
 CC SEQUENCE 2333 AA; 259645 MW; 4FC667DCC521BC60 CRC64;
 SQ SEQUENCE 2333 AA; 259645 MW; 4FC667DCC521BC60 CRC64;

Query Match 4.4%; Score 126.5; DB 1; Length 2333;
 Best Local Similarity 22.2%; Pred. No. 0.56;
 Matches 78; Conservative 41; Mismatches 144; Indels 89; Gaps 15;

QY 72 PGELIWSAPLGPDLNPLNYSILARMYNGVAGGEVQVILAGNAFTAGKIIFAAVPPNFPTE 131
 DB 360 PFGYLTLELPDTHRGVFGHLLVSAYMRNGDVEVSAGVNGQFNGGCLLVAMVPEKAFD 419
 QY 132 GLSPSQVTMFPHIIVDVRLQLEPVLPLPDVRNRYFY-HYNQSDSTIKLFIAMLYTPLRANN 190
 DB 420 TREKYQLTLFPHQFISPRNTNMTAHTVTPYLGNRVDQYKHKPWT--LVVMVLSPLTVSN 477

Wed Jun 2 09:13:30 2004

us-09-926-799-5.rsp

```
QY 191 AGDDVFTVSCRVLTRSPDFDFIFLVPTV-----ESRTKFTVPILTVEEMNSRPPIP 245
Db 478 TA-----APQIKVYANIAPTYVHVAGELPSKEGIFPVACADGYGGLVTTDP 523
QY 246 -----LEKLYTGPSSAFVQPONGRCCTDGVLLGTTTOLSAVNIC-TF----- 286
Db 524 KTADPVYGVYNNPKTNY-----PGRF-----TNLLDVAEACPTFLRFDGKPYVVT 570
QY 287 RGDVTHIAGSHDYTMNLASQNSNYDPTTEEIPAPLGTPDFVGKIQGMLTQTTREDGSTR 346
Db 571 RADDTRLLAKFD--VSLAKHMSN-----TYLSGIAQYVYQYSGTINL 611
QY 347 H-KATVSTGS-----VHFTPKLGSVQYTTDT-----NNDFOGQNTKFT 384
Db 612 HFMTGSTDSKARYMVAYIPP--GVETPPDTPEAAHCHIAEWDGTLANSKFT 661
```

Search completed: June 1, 2004, 13:47:16
Job time : 8.51141 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:36:02 ; Search time 10.4002 Seconds
(without alignments)
4985.230 Million cell updates/sec

Title: US-09-926-799-5
Perfect score: 2886
Sequence: 1 MKMASNDANPSDGTANLVP.....VNQFVTLAPMGAGRRRAL 539
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2804	97.2	539	2 S40111	capsid protein - h
2	1867.5	64.7	542	2 S60616	capsid protein - h
3	1161.5	40.2	530	2 B37471	capsid protein - N
4	1110	38.5	546	2 B37491	major capsid prote
5	337.5	11.7	576	2 A53982	capsid protein - E
6	329	11.4	2344	2 S64740	genome polyprotein
7	323	11.2	2344	1 RRWRH	genome polyprotein
8	319	11.1	2344	2 S55399	genome polyprotein
9	300.5	10.4	671	1 VCHWF9	coat protein - fel
10	296.5	10.3	702	1 A48562	coat protein - San
11	294	10.2	668	1 VCHWFC	coat protein - fel
12	293.5	10.2	668	1 JQ2354	capsid protein - f
13	293	10.2	668	1 VCHWFF	coat protein - fel
14	288.5	10.0	703	1 A48562	coat protein - San
15	288	10.0	668	2 JQ2356	capsid protein - f
16	133.5	4.6	2194	1 GNNYE7	genome polyprotein
17	130.5	4.5	1007	2 H81670	conserved hypotet
18	127	4.4	2332	1 GNNYF	genome polyprotein
19	126.5	4.4	2333	1 GNNY2F	genome polyprotein
20	126	4.4	948	2 T11678	hypothetical prote
21	123	4.3	26226	1 I38344	titin, cardiac mus
22	122.5	4.2	940	2 D89723	protein F39D8.1b l
23	122.5	4.2	945	2 T21998	hypothetical prote
24	122.5	4.2	2201	1 GNNYA9	genome polyprotein
25	122	4.2	1011	1 GNNYCI	genome polyprotein
26	121.5	4.2	733	2 JQ1892	capsid protein - f
27	121	4.2	3473	1 A46112	genome polyprotein
28	121	4.2	3473	2 S27927	polyprotein - rice
29	120.5	4.2	733	2 JQ1891	capsid protein - f

30	118.5	4.1	2332	1 GNNY4F	genome polyprotein
31	116	4.0	757	2 J80198	genome polyprotein
32	115	4.0	2130	2 AB0821	probable exported
33	115	4.0	2336	2 S37077	genome polyprotein
34	112.5	3.9	2175	1 GNNYBE	genome polyprotein
35	111.5	3.9	2233	2 B95075	beta-galactosidase
36	110	3.8	651	2 S53415	hypothetical prote
37	109.5	3.8	3890	2 C89921	hypothetical prote
38	109	3.8	782	2 I48746	semaphorin C - mou
39	108.5	3.8	1615	2 B49502	protein-tyrosine-p
40	108.5	3.8	1767	2 A49502	protein-tyrosine-p
41	108	3.7	13055	2 T16580	hypothetical prote
42	107.5	3.7	401	2 F90810	flagellar hook pro
43	107.5	3.7	401	2 B85670	flagellar biosynth
44	107	3.7	2481	2 D90011	FmB protein (impo
45	106.5	3.7	686	2 D96611	probable CRK1 prot

ALIGNMENTS

RESULT 1

S40111

capsid protein - human calicivirus (strain Bristol isolate B493)

C:Species: human calicivirus

A:Variety: strain Bristol isolate B493

C:Date: 25-Dec-1994 #sequence_revision 27-Feb-1997 #text_change 28-Jul-2000

OMASGSGSNTM S40111

K:Green, S.M.; Dingle, K.E.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.

submitted to the EMBL Data Library, December 1993

A:Description: Human enteric caliciviridae: a new prevalent SRSV group defined by RNA-d

A:Reference number: S40111

A:Accession: S40111

A:Molecule type: genomic RNA

A:Residues: 1-539 <GRE>

A:Cross-references: EMBL:X76716; NID:G436410; PID:CAA54134.1; PID:G436411

A:Experimental source: human enteric calicivirus strain Bristol isolate B493

C:Superfamily: human calicivirus capsid protein

C:Keywords: capsid protein; coat protein

Query Match	97.2%;	Score	2804;	DB 2;	Length	539;			
Best Local Similarity	96.3%;	Pred. No.	6.6e-200;						
Matches	519;	Conservative	13;	Mismatches	7;	Indels	0;	Gaps	0;
QY	1	MKMASNDANPSDGTANLVP	EVNNEVMALEPVVGAAIAAPVAGQONVIDPWIRN	FVQAP	60				
Db	1	MKMASNDANPSDGTANLVP	EVNNEVMALEPVVGAAIAAPVAGQONVIDPWIRN	FVQAP	60				
QY	61	GGEFTVSPRNAPGEILWSAPL	GPLDLPYLSHLARMYNGYAGGFVQVILAGNAFTAGKII	120					
Db	61	GGEFTVSPRNAPGEILWSAPL	GPLDLPYLSHLARMYNGYAGGFVQVILAGNAFTAGKVI	120					
QY	121	FAAAPPNPTEGLSPSQVTMPFPII	VDVROLEPVLIPDPVRNPFYHYNQSDNSTIKLIA	180					
Db	121	FAAAPPNPTEGLSPSQVTMPFPII	VDVROLEPVLIPDPVRNPFYHYNQANDSTIKLIA	180					
QY	181	MLYTPLRANAGDDVFTVSCRVL	TRPSPDFPFLVPPTVESRTKPFVTPILTVVEMSGNS	240					
Db	181	MLYTPLRANAGDDVFTVSCRVL	TRPSPDFPFLVPPTVESRTKPFVTPVLTVEMSGNS	240					
QY	241	RPPIPLEKLYTGPSSAFVVPQNG	RCTTGDVLLGTQLLSAVNICFTRPGDVTHIAGSHDYT	300					
Db	241	RPPIPLEKLYTGPSSAFVVPQNG	RCTTGDVLLGTQLLSAVNICFTRPGDVTHIAGSHDYT	300					
QY	301	MNLASQNSNDPTEIEIPALPGT	PDFVCKIQGLTQTTREDGSTRAHKATVSTGSHFTP	360					
Db	301	MNLASQNSNDPTEIEIPALPGT	PDFVCKIQGLLTQTTTRADGSTRAHKATVSTGSHFTP	360					
QY	361	KLGSVQYTTDNTDNFQTKFT	PVGVIQDGNHQNPEQQWVLPNYSGRGTGNVHLAPA	420					
Db	361	KLGSVQYTTDNTDNFQTKFT	PVGVIQDGHQHNPEQQWVLPNYSGRGTGNVHLAPA	420					
QY	421	VAPTFPGEQILLFRSTMPG	CSGYPNNMLDCLLPQEWVQHFQCEAAPAQSDVALLRFVNPD	480					

Db 421 VAPTFGEQLLFFRSTMPGCGYPNNMLDCLLPQEWLHFYQEAAPQAQSDVALLRFVNPDP 480

Qy 481 TGRVLPFCKLHKSGVYVTAHTGPHDLVPPNGYFRFDSWVQFYTLAPMGNGAGRRRAL 539

Db 481 TGRVLPFCKLHKSGVYVTAHTGPHDLVPPNGYFRFDSWVQFYTLAPMGNGAGRRRAL 539

RESULT 2

S60616

Capsid protein - human calicivirus (strain Melksham)

C:Species: human calicivirus

A:Variety: strain Melksham

C:Date: 23-May-1997 #sequence_revision 23-May-1997 #text_change 28-Jul-2000

C:Accession: S60616

R:Green, S.M.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.

Virus Res. 37, 271-283, 1995

A:Title: Capsid diversity in small round-structured viruses: molecular characterization

A:Reference number: S60615; MUID:96136658; PMID:8533462

A:Accession: S60616

A:Molecule type: Genomic RNA

A:Residues: 1-542 <GRE>

A:Cross-references: EMBL:X81879; NID:g976077; PIDN:CAA57462.1; PID:g976079

A:Experimental source: Strain Melksham

A:Note: it is uncertain whether Met-1 or Met-3 is the initiator

C:Superfamily: human calicivirus capsid protein

C:Keywords: capsid protein; coat protein

Query Match 64.7%; Score 1867.5; DB 2; Length 542;

Best Local Similarity 63.3%; Pred. No. 1.6e-130;

Matches 345; Conservative 79; Mismatches 108; Indels 13; Gaps 4;

Qy 1 MKMASNDANPSDGTANLVEPVNNEVMALEPVVGAIAAPVAGQOQNVDPWIRNPFVQAP 60

Db 1 MKMASNDAASTDGAAGLVPEVNNEVMALEPVVGAIAAPVGTQTIIDPWIRANFVQAP 60

Qy 61 GGEFTVSPNAPGEILWSAPLGPDLNPLYSLHARMYNGVAGGEVQVILAGNAFTAGKII 120

Db 61 NGEFTVSPNAPGEVLLNLELGPDLNPLYLAHARMYNGVAGGEVQVILAGNAFTAGKLV 120

Qy 121 FAAPVPNFPTEGLSPQVTMFPHIIIVDVRQLEPVLPLDPVRNFFHYNQSDNSTIKLIA 180

Db 121 FAAPVPNFPTEGLSPQVTMFPHIIIVDVRQLEPVLPLDPVRNFFHYNQSDNSTIKLIA 180

Qy 181 MLYTPLRNAGDDVFTVSCRVLTRSPDFDFLVPPTVESRTKPTFPIITVEEMSNS 240

Db 181 MLYTPLRNAGDDVFTVSCRVLTRSPDFDFLVPPTVESRTKPTFPIITVEEMSNS 240

Qy 241 RFPPILEKLYTGSSAFVQPNQGRCTTGGVLLGTTLQLSAVNICTFRGDVTHIAGSHDYT 300

Db 241 RFPPILEKLYTGSSAFVQPNQGRCTTGGVLLGTTLQLSAVNICTFRGDVTHIAGSHDYT 300

Qy 301 MN--LASQNSNYDPTBEIPAPLGTDFVGVKIQGMLTQTTRDQ-----STRAHKATVS 352

Db 301 NNVITNLNPSFPDSEDIAPLGPVDFQGRVFGVISORDKQNAAGHSEPANRGHDAVVP 360

Qy 353 TGSVHFTPKLGSVQYTTDNDFTQNTKFTFPGVIGVQDGNHNEPQOQVLPVNSGRTG 412

Db 361 TYTAQYTPKLGQIQIGTWTQTDLTWNQVFKFTPVGL--NDTEHFN---QWVVPRIYALN 415

Qy 413 HNVHLAPAVAPTFPGEQLLFFRSTMPGCGYPNNMLDCLLPQEWVQHFQCBAAQAQSDVA 472

Db 416 LNTNLAPSVAVPFGEQLLFFRSHPLPLKGGVGNFAIDCLLPQEWVQHFQCBAAQAQSDVA 475

Qy 473 LLRFVNPDTGRVLPFCKLHKSGVYVTAHTGPHDLVIPPNGYFRFDSWVQFYTLAPMGNG 532

Db 476 LVRYINPTDGRALFEAKLHRAAGMTVSSNTSAPVVPVANGYFRFDSWVQFYSLAPMGNG 535

Qy 533 AGRRR 537

Db 536 NGRRR 540

RESULT 3

B37471

Capsid protein - Norwalk virus

C:Species: Norwalk virus

C:Date: 24-Feb-1994 #sequence_revision 25-Apr-1997 #text_change 28-Jul-2000

C:Accession: B37471

R:Jiang, X.; Wang, M.; Wang, K.; Estes, M.K.

Virology 195, 51-61, 1993

A:Title: Sequence and genomic organization of Norwalk virus.

A:Reference number: A37471; MUID:93303939; PMID:8391187

A:Accession: B37471

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: genomic RNA

A:Residues: 1-530 <JIA>

A:Cross-references: GB:M87661; NID:g1061311; PIDN:AAB50466.1; PID:g1061313

A:Note: sequence extracted from NCBI backbone (NCBI:P134157)

C:Superfamily: human calicivirus capsid protein

Query Match 40.2%; Score 1161.5; DB 2; Length 530;

Best Local Similarity 45.4%; Pred. No. 3.1e-78;

Matches 250; Conservative 83; Mismatches 181; Indels 37; Gaps 14;

Qy 1 MKMASNDANPS-DGST--ANLVEPVN-NEVMALEPVVGAIAAPVAGQOQNVDPWIRNPF 56

Db 1 MMASKDATSDVDCASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIINNF 60

Qy 57 VOAPGGEFTVSPNAPGEILWSAPLGPDLNPLYSLHARMYNGVAGGEVQVILAGNAFTA 116

Db 61 VOAPGGEFTVSPNAPGEILWSAPLGPDLNPLYSLHARMYNGVAGGEVQVILAGNAFTA 120

Qy 117 GKIIFAAPVPNFPTEGLSPQVTMFPHIIIVDVRQLEPVLPLDPVRNFFHYNQSDNSTI 176

Db 121 GKIIIVSCIPPCFGSHNLITIAQATLPHVIADVRLDIEVPLEDVRNVLHNDRNQOTM 180

Qy 177 KLIAMLYTPLRANAGDDVFTVSCRVLTRSPDFDFLVPPTVESRTKPTFPIITVEE 236

Db 181 RLVCMLYTPLETTGGTGGDSFVYAGRVMTCPSPDFNLFVLPPTVEQKTRPPTLNLPLSS 240

Qy 237 MSNRFPIPLEKLYTGSSAFVQPNQGRCTTGGVLLGTTLQLSAVNICTFRGDVTHIAGS 296

Db 241 LNSRAPAPISSIGISPDNVQSVQFQNGRCLDGLVGTTPVLSLHVAKIRGT-----S 294

Qy 297 HDYTNLASQNSNYDPTBEIPAPLGTDFVGVKIQGMLTQTTRDQSTRAHKATVSTGVS 356

Db 295 NGTVNLINTELDTGTPHPPEG-PAPIGFPD-LGGCDWHINMT--QFGHSSQYQYDVTTPD 350

Qy 357 HFTPKLGSVQYTTDNDFTQNTKFTFPGVIGVQDGNHNEPQOQVLPVNSGRTG 412

Db 351 TEVPHLGSIQ-----ANGIGSNY-----VGVLWSIPSPSPSGSVDLWKIPNYGSSIT 400

Qy 413 HNVHLAPAVAPTFPGEQLLFFRSTMPGCGYPNNMLDCLLPQEWVQHFQCBAAQAQSDVA 472

Db 401 EATHLAPSVPYPPGFEVLVFFEMKMPGAY---NLPLLQEVYISHLAEQAQPTVGAA 457

Qy 473 LLRFVNPDTGRVLPFCKLHKSGVYVTAHTGPHDLVIPPNGYFRFDSWVQFYTLAP 528

Db 458 LLHYVDPDTGRNLGFKAYPDGFLCTVPNGASSGQQQL--PINGVVFVWSVRFYQLKP 515

Qy 529 MCGAGRRAL 539

Db 516 VGTASSARGRL 526

RESULT 4

B37491

major capsid protein [similarity] - Southampton virus

N:Alternate names: orf2 protein

C:Species: Southampton virus

C:Date: 22-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 18-Aug-2000

C:Accession: B37491

R:Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.

Science 259, 516-519, 1993

A:Title: Sequence and genome organization of a human small round-structured (Norwalk-like

C;Keywords: blocked amino end; polyprotein

Query Match 11.4%; Score 329; DB 2; Length 2344;
Best Local Similarity 25.3%; Pred. No. 1.3e-15;
Matches 137; Conservative 62; Mismatches 213; Indels 130; Gaps 23;

QY 4 ASNDANPSDGTANLVPEVNNVMALEPVVGAIAAPVAGQONVID---PWIRNNFVQAP 60
DB 1785 ASVPGTTTDDGMDPGVV--ATTSVTAENSSASIAATAGIGGPPQVDDQETW-RTNFYV-- 1839

QY 61 GGEFTVSPRNAPGEILMSAPLGDLPNLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
DB 1840 NDVFTWSVADAPGSILYTVQHSPPNNPFTAVLSQMYAGWAGGQFRFIVAGSGVFGRLV 1899

QY 121 FAAPVPNFTPEGLSPS-QVTMFPFHIIVDRQLEPVLPLPDVRRNPFHYNQSNDSITIKLI 179
DB 1900 RAVIIPPGIE---IGPGLVRQFPFHVVIDARSLSEPTITMPDLRPNMYHPTGDPGLVPTLV 1956

QY 180 AMLYTLRANAGDDVFTVSCRVLTRPSDPDEFILVPP---TVESRTKP--FTVPILTV 234
DB 1957 LSVYNNL-INPFGGTSAIQVTVETRPSEDFEFWMIRAPSSKTVDISIPAGLLTTPVLT- 2014

QY 235 EEMSNRRFPIELEKLYTGPSSAFVQPONG---RCTTDGVLGTTQ-LSAVNICTRGDV 290
DB 2015 -----GVGNDNRWG--QIVGLQVPVGGFSTCNRHWNLSGTVMSSPRF---GDI 2060

QY 291 THIAGSHDYTMNLAS---QNW-----SNYD-PTEETIPAPLGTDP----- 325
DB 2061 DHRRGSASYSGSNATNLQFWYANAGSAIDNPISQV-APDGFDMSPVFPNGGIPAAGW 2119

QY 326 -----FVGKIQGMLTOTTRDGSTRAHKA---TVSTGS--- 355
DB 2120 VGFCAIWNNSGAPNVTTVQAYELGFATGAPGNLQPTTNTSGAQTVAKSIYAVVTGTAQN 2179

QY 356 -----VHFTPKLGSVQYTTDTNNDFTQGTQKFTPVGVITQ----- 390
DB 2180 PAGLFVWASGIISTPANAITYTPQDRIIVTTPGTAAAPVGKNTPIIMEASVVRRTGDVN 2239

QY 391 -----DGNHNEPQQWVLPNYSRGTGNVHLAPAVAPTFPGEQLLFFRSTMPGCGYPNM 446
DB 2240 ATAGSANGTQGTGSQLPVTIGLSLNNYSSA-----LMPGQ---FFVWQLTFASGFMEI 2291

QY 447 NL 448
DB 2292 GL 2293

RESULT 8
S55399
genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)
A:Species: rabbit hemorrhagic disease virus
A:Variety: isolate BS89
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 18-Jun-1999
C:Accession: S55399
R:Rossi, C.
submitted to the EMBL Data Library, May 1995
A:Reference number: S55399
A:Accession: S55399
A:Molecule type: genomic RNA
A:Residues: 1-2344 <ROS>
A:Cross-references: EMBL:X87607; NID:g854640; PIDN:CAA60910.1; PID:g854641
A:Experimental source: isolate BS89
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: polyprotein

Query Match 11.1%; Score 319; DB 2; Length 2344;
Best Local Similarity 25.1%; Pred. No. 7e-15;
Matches 136; Conservative 64; Mismatches 213; Indels 128; Gaps 24;

QY 4 ASNDANPSDGTANLVPEVNNVMALEPVVGAIAAPVAGQONVID---PWIRNNFVQAP 60
DB 1785 ASVPGTTTDDGMDPGVV--ATTSVTAENSSASIAATAGIGGPPQVDDQETW-RTNFYV-- 1839

QY 61 GGEFTVSPRNAPGEILMSAPLGDLPNLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
DB 1840 NDVFTWSVADAPGSILYTVQHSPPNNPFTAVLSQMYAGWAGGQFRFIVAGSGVFGRLV 1899

QY 121 FAAPVPNFTPEGLSPS-QVTMFPFHIIVDRQLEPVLPLPDVRRNPFHYNQSNDSITIKLI 179
DB 1900 RAVIIPPGIE---IGPGLVRQFPFHVVIDARSLSEPTITMPDLRPNMYHPTGDPGLVPTLV 1956

QY 180 AMLYTLRANAGDDVFTVSCRVLTRPSDPDEFILVPP---TVESRTKP--FTVPILTV 234

C;Keywords: blocked amino end; polyprotein

Query Match 11.4%; Score 329; DB 2; Length 2344;
Best Local Similarity 25.3%; Pred. No. 1.3e-15;
Matches 137; Conservative 62; Mismatches 213; Indels 130; Gaps 23;

QY 4 ASNDANPSDGTANLVPEVNNVMALEPVVGAIAAPVAGQONVID---PWIRNNFVQAP 60
DB 1785 ASVPGTTTDDGMDPGVV--ATTSVTAENSSASIAATAGIGGPPQVDDQETW-RTNFYV-- 1839

QY 61 GGEFTVSPRNAPGEILMSAPLGDLPNLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
DB 1840 NDVFTWSVADAPGSILYTVQHSPPNNPFTAVLSQMYAGWAGGQFRFIVAGSGVFGRLV 1899

QY 121 FAAPVPNFTPEGLSPS-QVTMFPFHIIVDRQLEPVLPLPDVRRNPFHYNQSNDSITIKLI 179
DB 1900 RAVIIPPGIE---IGPGLVRQFPFHVVIDARSLSEPTITMPDLRPNMYHPTGDPGLVPTLV 1956

QY 180 AMLYTLRANAGDDVFTVSCRVLTRPSDPDEFILVPP---TVESRTKP--FTVPILTV 234
DB 1957 LSVYNNL-INPFGGTSAIQVTVETRPSEDFEFWMIRAPSSKTVDISIPAGLLTTPVLT- 2014

QY 235 EEMSNRRFPIELEKLYTGPSSAFVQPONG---RCTTDGVLGTTQ-LSAVNICTRGDV 290
DB 2015 -----GVGNDNRWG--QIVGLQVPVGGFSTCNRHWNLSGTVMSSPRF---GDI 2060

QY 291 THIAGSHDYTMNLAS---QNW-----SNYD-PTEETIPAPLGTDP----- 325
DB 2061 DHRRGSASYSGSNATNLQFWYANAGSAIDNPISQV-APDGFDMSPVFPNGGIPAAGW 2119

QY 326 -----FVGKIQGMLTOTTRDGSTRAHKA---TVSTGS--- 348
DB 2120 VGFCAIWNNSGAPNVTTVQAYELGFATGAPGNLQPTTNTSGAQTVAKSIYAVVTGTAQN 2179

QY 349 -----ATVSTGSHVFTPKLGSVQYTTDTNNDFTQGTQKFTPVGVITQ----- 390
DB 2180 PAGLFVWASGIISTPANAITYTPQDRIIVTTPGTAAAPVGKNTPIIMEASVVRRTGDVN 2239

QY 391 -----DGNHNEPQQWVLPNYSRGTGNVHLAPAVAPTFPGEQLLFFRSTMPGCGYPNM 446
DB 2240 ATAGSANGTQGTGSQLPVTIGLSLNNYSSA-----LMPGQ---FFVWQLTFASGFMEI 2291

QY 447 NL 448
DB 2292 GL 2293

RESULT 7
RRWRH
genome polyprotein - rabbit hemorrhagic disease virus
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C:Species: rabbit hemorrhagic disease virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 29-May-1998
C:Accession: A41039
R:Meyer, G.; Wirblich, C.; Thiel, H.J.
Virology 184, 664-675, 1991
A:Title: Rabbit hemorrhagic disease virus-molecular cloning and nucleotide sequencing
A:Reference number: A41039; MUID:91361557; PMID:1840711
A:Accession: A41039
A:Molecule type: Genomic RNA
A:Residues: 1-2344 <MEY>
A:Cross-references: GB:M67473
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: nucleotidyltransferase

Query Match 11.2%; Score 323; DB 1; Length 2344;
Best Local Similarity 25.1%; Pred. No. 3.6e-15;
Matches 136; Conservative 66; Mismatches 210; Indels 130; Gaps 24;

QY 4 ASNDANPSDGTANLVPEVNNVMALEPVVGAIAAPVAGQONVID---PWIRNNFVQAP 60
DB 1785 ASVPGTTTDDGMDPGVV--ATTSVTAENSSASIAATAGIGGPPQVDDQETW-RTNFYV-- 1839

Db 1957 LSVYNNL-INPFGSTSAIQVTVETRPSEDFEFVIMIRAPSSKTVDSISPAGLTTTPVLT- 2014
QY 235 EEMNSRFPPIPLEKLYTPSSAFVQPNQNG---RCTTDGVLGTTQ-LSAVNICTFRGDV 290
Db 2015 -----GVGNDNRWG--QIVGLQVPVGGFSTCNRHNLNGSTVGWSSPRF-----ADI 2060
QY 291 THAGSHDYTNLAS---QNW-SNYDPTBEIP-----APLGTPD----- 325
Db 2061 DHRGASYPGSGNATNVLFQWYANAGSAVDNPIQVAEPDGFPMGFPVPPNGPGIPAAAGMV 2120
QY 326 -----FVGKIQGMLTOTTRDGDSTRAHKA--TVSTGS----- 355
Db 2121 GFCAIWNSSGAPNVTTVQAYELGATGAPGNLQPTNTSGAQTVAKSIYAVVTGAQNP 2180
QY 356 -----VHFTPKLGSVQYTTDTNNDFQT-----GQNTKFTPVGVLIQ----- 390
Db 2181 AGLFVWASGVISTPNAIAITYTPQPDRIVTTTGTTPAAAPVGNKTPIMFASVVRRTGDVNA 2240
QY 391 ----DGNHQNEPQQWVLPNYSRGTGHNHVLAPAVAPTFFGEOLLFFRSTMPCCSGYPNNV 447
Db 2241 TAGSANGTQYGTGSOPLPVTIIGLSLNYSSA-----LMBGQ---FFVMQLTPASGFMEIG 2292
QY 448 L 448
Db 2293 L 2293
RESULT 9
VCWVF9
coat protein - feline calicivirus (strain P9)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C>Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text_change 28-Jul-2000
C:Accession: B43382; C45538; P00407; S23702
R:Carter, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.
Virology 190, 443-448, 1992
A:Title: Identification and sequence determination of the capsid protein gene of feline
A:Reference number: A45538; MUID:92117861; PMID:1731695
A:Accession: C45538
A:Molecule type: genomic RNA
A:Residues: 1-671 <CAR1>
A:Cross-references: GB:M86379; NID:G323877; PIDN:AAAT9327.1; PID:G323879
R:Carter, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.
Arch. Virol. 122, 223-235, 1992
A:Title: Identification and sequence determination of the capsid protein gene of feline
A:Reference number: A45538; MUID:92117861; PMID:1731695
A:Accession: C45538
A:Molecule type: genomic RNA; protein
A:Residues: 1-671 <CAR2>
A:Cross-references: GB:M86379; NID:G323877; PIDN:AAAT9327.1; PID:G323879
A:Experimental source: strain F9
A>Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)
R:Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.
J. Gen. Virol. 73, 2429-2433, 1992
A:Title: The cloning, sequencing and expression of a major antigenic region from the fel
A:Reference number: P00407; MUID:93019069; PMID:1402818
A:Accession: P00407
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 217-266, 'SI', 269-336, 'V', 338-395, 'S', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>
A:Cross-references: PIDN:AAAT9327.1; PID:G257083
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 10.4%; Score 300.5; DB 1; Length 671;
Best Local Similarity 25.3%; Pred. No. 2.7e-14;
Matches 141; Conservative 80; Mismatches 238; Indels 99; Gaps 28;
QY 12 DGSANLVPVNEV--MALEPVVGNATAAPVAGQNVDPWIRNPFVQAPGERTV--- 66
Db 127 DGSIT--APEQGTWGGVIAEBSAQMSSTAADMATGKSDSEW-----EAPFSFHTSVNW 178

QY 67 SPRNAPGEILMSAPLGPDLNPYLSHARMYAGGFEVQVILAGNATAGKIIIFAAVPP 126
Db 179 STSETQGGKILFKQSLGPLNPLEHLAKLYVAMSGSIEVRFSISGSGVFGGKLAIVVPP 238
QY 127 NFPTTEGLSPSOVT--MEPHIIVDVRLPEVLPLPDPVNRNPFYHYNQNSNDSTIKLIAMLY 183
Db 239 -----GVDPVQSTSMQYPHVLFDAQVPEPIFCLPDLRSLYHLMSDSDTT-SLIVIMY 292
QY 184 TPLRANAGDDVFTVSC--RVLTRPSDDPFIFLVPPTVEGRKTFPTVPILTVFEMSNSR 241
Db 293 NDL-INPYANDANSGCCIVTVETKPGDPFKHLLKPP--GSMLTRHGSIPSDLIPTKTS-- 347
QY 242 FPIPLEKLYTPSSAFVQPNQNG---QNGR-----CTTDGVLGTTQLSAVNICTFRGDVTHI 293
Db 348 --LWIGNRYSDITDFVIRPFVFOANRHDFENQETAGWSTPRFRISVTITEQNGAKLGI 405
QY 294 AGSHDYTNLASQWNSYD-PTEEIPAPLGTDFVGIQGMLTQTTREDGSGFRA--HKAT 350
Db 406 GVATDIYVGPDPDGMPTTIFGELIPA-----GDYA-----ITNGTGNDITATGYDTAD 455
QY 351 VSTGSVHF-----TPKLGSVQYTT-----DTNNDPQTGQNTKFTPVGVLIQ 390
Db 456 IIKNTNFRGMVYCGSLQRAWGDKKISNTAFITTATLGDNNKINPCNTIDQSKIIVWFQ 515
QY 391 D---GNHQNPEQQWVLPNYSG-----RTGHNHVLAPAVAPTFF-----GEOLLFFRSTM 437
Db 516 DNHVGKKAQTSDDTLALLGYTGIGQAIGSDRDRVVRIS-TLPETGARGGNHPIFYKNSI 574
QY 438 PGCSGYPNNMLDCLLPQEMVQHFQCEAA-----PAQSDVALLRFVNPDPTRVLECKLH 491
Db 575 K--LGIVIRSDIVFNSQ--ILHTRQSLNHYLLPDS--FAYRII---DSNGSWFDIGID 627
QY 492 KSGYVTVTAHTGPHDIVIP 509
Db 628 SDGFSFGVSGGKLEFP 645
RESULT 10
A48562
coat protein - San Miguel sea lion virus (serotype 1)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: A48562
R:Neill, J.D.
Virus Res. 24, 211-222, 1992
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel
eins.
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Accession: A48562
A:Molecule type: genomic RNA
A:Residues: 1-702 <NEI>
A:Cross-references: GB:M87481; NID:G334882; PIDN:AAAL6217.1; PID:G334884
A>Note: sequence extracted from NCBI backbone (NCBIN:113564, NCBIP:113565)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 10.3%; Score 296.5; DB 1; Length 702;
Best Local Similarity 26.1%; Pred. No. 5.7e-14;
Matches 141; Conservative 72; Mismatches 199; Indels 129; Gaps 30;
QY 11 SDG-STANLVPVNEVNALEBPVVG-----AAIAAPVAGQNVDPWIRNPFVQAPGERT 65
Db 153 SDGPGGADIVTEEQGTVVQQQPVPAQSALTTLAAASTGK--TVD-----CEWT 198
QY 66 V-----SPRNAPGEILMSAPLGPDLNPYLSHARMYAGGFEVQVILAGNAT 115
Db 199 TFFSYHTAVNNTTEAQKILFSLRALSPELNPYLRHISLSYSTWSGGIDVRFVTSGSGVF 258
QY 116 AKGIIFAAVPPNF-PTEGLSPSOVTMFPHIIIVDRLPEVLPLPDPVNRNPFYHYNQNSD 174
Db 259 GKKLAALIVPPGIEBEV--SPTML-QYPHVLFDARQTEFVFTIPIPIRKTLYHSMDDTDT 315

C:Accession: A40507; B40507; T09246
R:Neill, J.D.; Reardon, I.M.; Heinrikson, R.L.
J. Virol. 65, 5440-5447, 1991
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline calicivirus
A:Reference number: A40507; MUID:91374597; PMID:1716692
A:Accession: A40507
A:Molecule type: genomic RNA
A:Residues: 1-668 <NE1>
A:Cross-references: GB:M32819; NID:G323874; PIDN:AAA42925.1; PID:G323875
A:Accession: B40507
A:Molecule type: protein
A:Residues: 373-379; 403-419; 481-489; 560-566 <NE2>
R:Neill, J.D.
submitted to the EMBL Data Library, April 1998
A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.
A:Reference number: Z16626
A:Accession: T09246
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: genomic RNA
A:Residues: 1-668 <NE3>
A:Cross-references: EMBL:U13992; NID:G3056875; PIDN:AAC13993.1; PID:G537256
A:Experimental source: strain CFI/68 FIV
A:Superfamily: feline calicivirus coat protein
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.2%; Score 293; DB 1; Length 668;
Best Local Similarity 25.3%; Pred. No. 9.7e-14;
Matches 117; Conservative 66; Mismatches 195; Indels 84; Gaps 18;

QY 12 DGSITANLVPVNEV--MALEPVGAAIAAPVAGQNVDPWIRNFFVQAPGGEFTV--- 66
DB 127 DGSIT--TPEQGTWVGVIAPNAQMSAADMATKSVDSW-----EAFSFTSNV 178
QY 67 SPRNAPGELWSAPGLDNLNLYSLARMYNGYAGFEVQVILAGNAFTAGKIIFAAVPP 126
DB 179 STSTQKGLFKQSLGSLPLNLYLHLAKYVAGSVDVRFSGSGVFGGKLAIVVPP 238
QY 127 NFPTGLSPSQVT--MFPHIIVDVRQLEPVLIPDPVRNFFHYNQSNDSITKLIAMLY 183
DB 239 -----GIDPVQSTMLQYPHVLFDAQVEPFIPLASTLYHLMSDDTT--SLVIMY 292
QY 184 TPLRANNAGDVFTVSC--RVLTRSPDPDFLPLVPTVESRTKFTVPILTVVEMSNR 241
DB 293 NDL--INPYANDSNSSGIVTVETKPGDPKFLHLLKPP--GSMLTRGSIPLDKSS-- 347
QY 242 PIPLEKLYTGSSAFVQVQ--QNGR-----CTTDGVLGTTQLSANNICTFRGDVTHI 293
DB 348 --LWIGNRFWSIDITDFVIRPFVQANRHFDFNOETAGWSTPRFRPITITISVKESAKLGI 405
QY 294 AGSHDYTMNLASQNSNYDPTBEIP---APLGTDPFVGKIQGLMTQTTREDGSTR----- 345
DB 406 GVATDYIVGIPDGW---PDTTIFGELVPVGDYAITNGTNDITTAQYDAATERNNT 461
QY 346 -----AHKATVSTGSVHFTPKLGS-----VOYTTDNNDQ 376
DB 462 NFRGMVYICGSLQAWGDKKISNTAFITTGTVGAKLIPNSTIDQTKIAVFQDTHANKHVQ 521
QY 377 TQONT-----KFTPVGVIQGNHQNHPQWVLPNYSGRTHN 414
DB 522 TSDDTLALLGYTGIGEEATGADRDRVRVRSIVLPERGARGGNH 563

RESULT 14
C48562
coat protein - San Miguel sea lion virus (serotype 4)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
A:Accession: C48562
R:Neill, J.D.
Virus Res. 24, 211-222, 1992
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel

eins.
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Accession: C48562
A:Molecule type: genomic RNA
A:Residues: 1-703 <NE1>
A:Cross-references: GB:M87482; NID:G334886; PIDN:AAA16220.1; PID:G334888
A>Note: sequence extracted from NCBI backbone (NCBIP:1113567)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.0%; Score 288.5; DB 1; Length 703;
Best Local Similarity 27.3%; Pred. No. 2.2e-13;
Matches 101; Conservative 56; Mismatches 164; Indels 49; Gaps 16;

QY 11 SDG--STANLVPVNEVMALEPV---VGAATAAPVAGQNVDPWIRNFFVQAPGGEFTV 66
DB 153 SDGPGSAEIVTEOQTVVQQQPAPAPATLATLATATGKSVQEQEMTFFSYHTSNWSTV 212
QY 67 SPRNAPGELWSAPGLDNLNLYSLARMYNGYAGFEVQVILAGNAFTAGKIIFAAVPP 126
DB 213 ---ESQKILYQALNPSINPYLDHIAKLYSTWSGIDVRFVSGSGVFGGKLAALLVPP 269
QY 127 NF--PTEGLSPSQVTMFPHIIVDVRQLEPVLIPDPVRNFFHYNQSNDSITKLIAMLYTP 185
DB 270 GVEPIESVSMLOQ---YPHVLFDAQTEPVIFITPDIRKTLFHSMDTDTT--KLVINPY-- 323
QY 186 LRANNAGDVFTVSCRVLTFRSPDPDFLPLVPP--TVESRTKPF--TVPLTVEEMSRSR 242
DB 324 ---ENGVENKTCSTIVETRPSADFTFALLKPPGSLIKHSGIPSDLI PRNSAHWMGCR-- 378
QY 243 PIPLEKLYTGSSAFVQVQ-----NGRCTTQGVLLGTTQLSANNI--CTFRGDVTHIAGSH 297
DB 379 -----WMTISGFSVQPRVQSNRHFDFSTTTGWSTPYVYVPIEIKIQKV----GSN 427
QY 298 DYTMLASQNSNYDPTBEIPAPLGTDPFVGKIQGLMTQTTREDGSTRAKHAKTVSTGSHV 357
DB 428 N-----KMFHVITDALKVP--GIPD--GWPDITPDEKATNGNFSYSGSYRAGSTT 476
QY 358 FTPKLGSVQY 367
DB 477 IKPENSTHP 486

RESULT 15
J02356
capsid protein - feline calicivirus (strain KCD)
C:Species: feline calicivirus
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: J02356
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable
A:Reference number: J02354; MUID:94065683; PMID:7504075
A:Accession: J02356
A:Molecule type: mRNA
A:Residues: 1-668 <SEA>
A:Cross-references: GB:L09719; NID:G305107; PIDN:AAA16487.1; PID:G305108
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein

Query Match 10.0%; Score 288; DB 2; Length 668;
Best Local Similarity 26.9%; Pred. No. 2.3e-13;
Matches 118; Conservative 54; Mismatches 196; Indels 70; Gaps 18;

QY 12 DGSITANLVPVNEV--MALEPVGAAIAAPVAGQNVDPWIRNFFVQAPGGEFTV--- 66
DB 127 DGSIT--VTPEQGTWVGVIAPNAQMSAADMATKSVDSW-----EAFSFTSNV 178
QY 67 SPRNAPGELWSAPGLDNLNLYSLARMYNGYAGFEVQVILAGNAFTAGKIIFAAVPP 126
DB 179 STSTQKGLFKQSLGSLPLNLYLHLAKYVAGSIEVRFSISGSGVFGGKLAIVVPP 238

Result No.	Query No.	Score	Query		Length	DB	ID	Description
			Match	\$				
1	1	2819	100.0		530	12	Q9T39	Q9T39 norwalk-lik
2	2	2789	98.9		530	12	Q8384	Q8384 norwalk vir
3	3	2737	97.1		517	12	Q9J41	Q9J41 norwalk-lik
4	4	2713	96.2		530	12	Q8376	Q8376 norwalk vir
5	5	2093	74.2		544	12	Q9DU6	Q9DU6 chiba virus
6	6	2093	74.2		544	12	Q9Q7E	Q9Q7E chiba virus
7	7	2084	73.9		544	12	Q9118	Q9118 human calic
8	8	2081	73.8		544	12	Q918C	Q918C norwalk-lik
9	9	2076	73.6		544	12	Q91E3	Q91E3 human calic
10	10	2063	73.2		544	12	Q918D	Q918D norwalk-lik
11	11	2046	72.6		544	12	Q91V4	Q91V4 human calic
12	12	2044.5	72.5		540	12	Q9Y51	Q9Y51 norwalk vir
13	13	2044	72.5		544	12	Q9V36	Q9V36 human calic
14	14	2031	72.0		531	12	Q8V73	Q8V73 norwalk-lik
15	15	2029.5	72.0		546	12	Q8JW4	Q8JW4 norwalk-lik
16	16	2027.5	71.9		540	12	Q8BCA	Q8BCA5 human calic

```
QY 121 GKIIIVSCIPPGSGSHNLTAQAATLFPFHVIADVRTLDPIEVLDPVNRNVLFHNNDRNQTM 180
DB 121 GKIIIVSCIPPGSGSHNLTAQAATLFPFHVIADVRTLDPIEVLDPVNRNVLFHNNDRNQTM 180
QY 181 RLVCMLYTLPLRTGGTGDSEFVVGAVRMTCPSPDENFLFVPTVBQKTRPFTLNLPLSS 240
DB 181 RLVCMLYTLPLRTGGTGDSEFVVGAVRMTCPSPDENFLFVPTVBQKTRPFTLNLPLSS 240
QY 241 LNSRAPLPIISGMGISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVIN 300
DB 241 LNSRAPLPIISGMGISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVIN 300
QY 301 LTELDTGTPPHFPGAPIGFDPDLGGCDWHINMTQFHSSTQYDVTDPFVPHLGSIQ 360
DB 301 LTELDTGTPPHFPGAPIGFDPDLGGCDWHINMTQFHSSTQYDVTDPFVPHLGSIQ 360
QY 361 ANIGSGNYIGVLSWSPSPSHSGSOVDLWKIPNYGSSITEATHLAPSVPFGFGEVLVF 420
DB 361 ANIGSGNYIGVLSWSPSPSHSGSOVDLWKIPNYGSSITEATHLAPSVPFGFGEVLVF 420
QY 421 FMSKIPGPGAYSILPCLLPQEIYISHLASEQAPTVEAALLHYVDPDTGRTLGEFKAYPDGF 480
DB 421 FMSKIPGPGAYSILPCLLPQEIYISHLASEQAPTVEAALLHYVDPDTGRTLGEFKAYPDGF 480
QY 481 LTCVPNGASSGPOOLPINGVFVSVWSRFYQLKPVGTASSARGRLGLRR 530
DB 481 LTCVPNGASSGPOOLPINGVFVSVWSRFYQLKPVGTASSARGRLGLRR 530
RESULT 2
Q83884 PRELIMINARY; PRT; 530 AA.
AC Q83884;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 58 kDa capsid protein.
OS Norwalk virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=11983;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9303935; PubMed=8391187;
RA Jiang X., Wang M., Wang K., Estes M.K.;
RT "Sequence and genomic organization of Norwalk virus,";
RL Virology 195:51-61(1993).
DR EMBL: M87661; AAB50466.1; -.
DR PIR: B37471; B37471.
DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF00915; Calici_coat; 1.
SQ SEQUENCE 530 AA; 56571 MW; CREDB8D6366CD236 CRC64;
Query Match 98.9%; Score 2789; DB 12; Length 530;
Best Local Similarity 98.5%; Pred. No. 4.1e-219;
Matches 522; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY 1 MMASKDATSVDGASGAGQLVPEVNASDPLAMDPAVGSSTAVATAGVNPIDPWIINNF 60
DB 1 MMASKDATSVDGASGAGQLVPEVNASDPLAMDPAVGSSTAVATAGVNPIDPWIINNF 60
QY 61 VQAPQGETIISPNTPGCVLFDLSLGLPHLNFLLHLSQWYNGVGNMRVIRIMLAGNAFTA 120
DB 61 VQAPQGETIISPNTPGCVLFDLSLGLPHLNFLLHLSQWYNGVGNMRVIRIMLAGNAFTA 120
QY 121 GKIIIVSCIPPGSGSHNLTAQAATLFPFHVIADVRTLDPIEVLDPVNRNVLFHNNDRNQTM 180
DB 121 GKIIIVSCIPPGSGSHNLTAQAATLFPFHVIADVRTLDPIEVLDPVNRNVLFHNNDRNQTM 180
QY 181 RLVCMLYTLPLRTGGTGDSEFVVGAVRMTCPSPDENFLFVPTVBQKTRPFTLNLPLSS 240
DB 181 RLVCMLYTLPLRTGGTGDSEFVVGAVRMTCPSPDENFLFVPTVBQKTRPFTLNLPLSS 240
```

```
QY 241 LNSRAPLPIISGMGISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVIN 300
DB 241 LNSRAPLPIISGMGISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVIN 300
QY 301 LTELDTGTPPHFPGAPIGFDPDLGGCDWHINMTQFHSSTQYDVTDPFVPHLGSIQ 360
DB 301 LTELDTGTPPHFPGAPIGFDPDLGGCDWHINMTQFHSSTQYDVTDPFVPHLGSIQ 360
QY 361 ANIGSGNYIGVLSWSPSPSHSGSOVDLWKIPNYGSSITEATHLAPSVPFGFGEVLVF 420
DB 361 ANIGSGNYIGVLSWSPSPSHSGSOVDLWKIPNYGSSITEATHLAPSVPFGFGEVLVF 420
QY 421 FMSKIPGPGAYSILPCLLPQEIYISHLASEQAPTVEAALLHYVDPDTGRTLGEFKAYPDGF 480
DB 421 FMSKIPGPGAYSILPCLLPQEIYISHLASEQAPTVEAALLHYVDPDTGRTLGEFKAYPDGF 480
QY 481 LTCVPNGASSGPOOLPINGVFVSVWSRFYQLKPVGTASSARGRLGLRR 530
DB 481 LTCVPNGASSGPOOLPINGVFVSVWSRFYQLKPVGTASSARGRLGLRR 530
RESULT 3
Q9JH41 PRELIMINARY; PRT; 517 AA.
AC Q9JH41;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein (fragment).
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20150145; PubMed=10686033;
RA Yoda T., Terano Y., Shimada A., Suzuki Y., Yamazaki K., Sakon N.,
RA Oishi I., Utagawa E.T., Okuno Y., Shibata T.;
RT "Expression of recombinant Norwalk-like virus capsid proteins using a
bacterial system and the development of its immunologic detection.";
RL J. Med. Virol. 60:475-481(2000).
DR EMBL: AB028247; BAA89031.2; -.
DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF00915; Calici_coat; 1.
FT NON_TER 1
FT NON_TER 517
SQ SEQUENCE 517 AA; 55011 MW; BAD7DF35B1DD0DF6 CRC64;
Query Match 97.1%; Score 2737; DB 12; Length 517;
Best Local Similarity 99.2%; Pred. No. 7e-215;
Matches 513; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 8 ATSSVDGASGAGQLVPEVNASDPLAMDPAVGSSTAVATAGVNPIDPWIINNFVQAPQGE 67
DB 1 ATSSVDGASGAGQLVPEVNTSDPLAMDPAVGSSTAVATAGVNPIDPWIINNFVQAPQGE 60
QY 68 FTISPNTPGCVLFDLSLGLPHLNFLLHLSQWYNGVGNMRVIRIMLAGNAFTAGKIIVSC 127
DB 61 FTISPNTPGCVLFDLSLGLPHLNFLLHLSQWYNGVGNMRVIRIMLAGNAFTAGKIIVSC 120
QY 128 IPPFGSGSHNLTAQAATLFPFHVIADVRTLDPIEVLDPVNRNVLFHNNDRNQTMRLVCMY 187
DB 121 IPPFGSGSHNLTAQAATLFPFHVIADVRTLDPIEVLDPVNRNVLFHNNDRNQTMRLVCMY 180
QY 188 TPLRTGGTGDSEFVVGAVRMTCPSPDENFLFVPTVBQKTRPFTLNLPLSSLSNSRAP 247
DB 181 TPLRTGGTGDSEFVVGAVRMTCPSPDENFLFVPTVBQKTRPFTLNLPLSSLSNSRAP 240
QY 248 LPISGMGISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVINLDELDT 307
DB LPISGMGISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVINLDELDT 307
```



```
Db 241 LPISGMGISPNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGTSNGTVINLTDLGT 300
Qy 308 PFHPEGPAPICFPDLGGCDWHNNMTQFCHSSQTYDYDTPDTFVPHLGSIQANGISG 367
Db 301 PFHPEGPAPICFPDLGGCDWHNNMTQFCHSSQTYDYDTPDTFVPHLGSIQANGISG 360
Qy 368 NYIGVLSWVSPSPSPSHPSGSQVDLWKIPNYGSSITEATHLAPSVYPPGFEVLVFPMSKIPG 427
Db 361 NYIGVLSWVSPSPSPSHPSGSQVDLWKIPNYGSSITEATHLAPSVYPPGFEVLVFPMSKIPG 420
Qy 428 PGAYSLPCLLPQEYISHLASQAPVGEAALLHYVDPTGRTLGFEKAYPDGFLTCVPNG 487
Db 421 PGAYSLPCLLPQEYISHLASQAPVGEAALLHYVDPTGRTLGFEKAYPDGFLTCVPNG 480
Qy 488 ASSGQQQLPINGVFVFSWVSRRFYQLKPVGTASSARG 524
Db 481 ASSGQQQLPINGVFVFSWVSRRFYQLKPVGTASTARG 517

RESULT 4
ID Q83876 PRELIMINARY; PRT; 530 AA.
AC Q83876;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Norwalk virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=11983;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SRSV-KY-89/89/J;
EX MEDLINE=94335115; PubMed=8057474;
RA Wang J., Jiang X., Madore H.P., Gray J., Desselberger U., Ando T.,
RA Sato Y., Oishi I., Lew J.F., Green K.Y., et al.;
RT "Sequence characterization of small, round-structured viruses in the Norwalk
RT virus group."
RL J. Virol. 68:6590-6596(1994).
DR EMBL; L23828; AA59229.1;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 530 AA; 56402 MW; 9C0E2C3199CCF051 CRC64;

Query Match 96.2%; Score 2713; DB 12; Length 530;
Best Local Similarity 96.8%; Pred. No. 6.6e-213;
Matches 513; Conservative 2; Mismatches 15; Indels 0; Gaps 0;

Qy 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIIINNF 60
Db 1 MMASKDATSSVDGASASQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIIINNF 60
Qy 61 VQAPQGEFTISPNPTPGVLDLSGLPHLNPFLHLSQMYNGWGNMVRIMLAGNAFTA 120
Db 61 VQAPQGEFTISPNPTPGVLDLSGLPHLNPFLHLSQMYNGWGNMVRIMLAGNAFTA 120
Qy 121 GKIIIVSCIPPGGSHNLITIAQATLPFHVIADVRTLDPIEVLVDVRLVFNHNDNQTM 180
Db 121 GKIIIVSCIPPGGSHNLITIAQATLPFHVIADVRTLDPIEVLVDVRLVFNHNDNQTM 180
Qy 181 RLVCMLYTLRTGGTGSDSFVAGVMTCPSPDFNLFVPTVEQKTRPFTLPNPLSS 240
Db 181 RLVCMLYTLRTGGTGSDSFVAGVMTCPSPDFNLFVPTVEQKTRPFTLPNPLSS 240
Qy 241 LSNRAPLPISGMGISPNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGTSNGTVIN 300
Db 241 LSNRAPLPISGMGISPNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGTSNGTVIN 300
Qy 301 LTDLGTTFHPEGPAPICFPDLGGCDWHNNMTQFCHSSQTYDYDTPDTFVPHLGSIQ 360
Db 301 LTDLGTTFHPEGPAPICFPDLGGCDWHNNMTQFCHSSQTYDYDTPDTFVPHLGSIQ 360
```

```
Qy 361 ANGISGNYIGVLSWVSPSPSHPSGSQVDLWKIPNYGSSITEATHLAPSVYPPGFEVLVFP 420
Db 361 ANGISGNYIGVLSWVSPSPSHPSGSQVDLWKIPNYGSSITEATHLAPSVYPPGFEVLVFP 420
Qy 421 FMSKIPGAYSLPCLLPQEYISHLASQAPVGEAALLHYVDPTGRTLGFEKAYPDGFL 480
Db 421 FMSKIPGAYSLPCLLPQEYISHLASQAPVGEAALLHYVDPTGRTLGFEKAYPDGFL 480
Qy 481 LTCVPNGASSGQQQLPINGVFVFSWVSRRFYQLKPVGTASSARGRLGLRR 530
Db 481 LTCVPNGASSGQQQLPINGVFVFSWVSRRFYQLKPVGTASTARGRLGLRR 530
```

RESULT 5

```
Qy Q9DU46 PRELIMINARY; PRT; 544 AA.
Db Q9DU46;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Chiba virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=99565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Chiba 407/1987/JP;
RA Someya Y., Takeda N., Miyamura T.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Chiba 407/1987/JP;
EX MEDLINE=20569531; PubMed=11118371;
RA Someya Y., Takeda N., Miyamura T.;
RT "Complete nucleotide sequence of the Chiba virus genome and functional
RT expression of the 3C-like protease in Escherichia coli."
RL Virology 278:490-500(2000).
DR EMBL; AB042808; BAB18267.1;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58351 MW; 1DB7A4E81FBB246F CRC64;
```

Query Match 74.2%; Score 2093; DB 12; Length 544;
Best Local Similarity 72.7%; Pred. No. 3e-162;
Matches 396; Conservative 53; Mismatches 80; Indels 16; Gaps 8;

```
Qy 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIIINNF 60
Db 1 MMASKDATSPSADGATGAGQLVPEVNTADPIIDPVAGSSTALATAGQVNLIDPWIIINNF 60
Qy 61 VQAPQGEFTISPNPTPGVLDLSGLPHLNPFLHLSQMYNGWGNMVRIMLAGNAFTA 120
Db 61 VQAPQGEFTISPNPTPGVLDLSGLPHLNPFLHLSQMYNGWGNMVRIMLAGNAFTA 120
Qy 121 GKIIIVSCIPPGGSHNLITIAQATLPFHVIADVRTLDPIEVLVDVRLVFNHNDNQTM 180
Db 121 GKIIIVSCIPPGGSHNLITIAQATLPFHVIADVRTLDPIEVLVDVRLVFNHNDNQTM 180
Qy 181 RLVCMLYTLRTGGTGSDSFVAGVMTCPSPDFNLFVPTVEQKTRPFTLPNPL 238
Db 180 RLVCMLYTLRTGGTGSDSFVAGVMTCPSPDFNLFVPTVEQKTRPFTVPNIPL 239
Qy 239 SSLNSRAPLPISGMGISPNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRG--TNG 296
Db 240 KYLSNSRIPNPIEGMSLSPDOTQNVQFQNGRCTLGRLVGTTPVSVLSOLCKFRGRTSQ 299
Qy 297 TVINLTDLGTTFHPEGPAPICFPDLGGCDWHNNMTQFCHSSQTYDYDTPDTFVPHLGSIQ 351
Db 300 RVLNLTDLGSGFFMAFAAPAPAGPDLGSCDWHNNMTQFCHSSQTYDYDTPDTFVPHLGSIQ 359
```

```

QY 352 FVPHLGSII--QANGIGSGNYIGVLWSVSPSHPSGSQVDLWKIPNYGSSITEATHLAPSV 409
Db 360 FVPHLSSITLDENVSSGGDIYIGTQWTSPPSDGGANTNFWKIPDYGSSLAESQLAPAV 419
QY 410 YPGGGEVLVFFMSKIPGP--GAYSL-PCLLPQEIYISHLASEQAPTVEAALLHYVDPD 465
Db 420 YPGGNEVIIVFMASIPGPNQSGSNLVPCLLPQEIYTHFISEQAPIQGEAALLHYVDPD 479
QY 466 TGRTLGEFKAYPDGFLTCVPNGASSGPQQLPINGVVFVSWVSRFYQLKPVGTASSARGR 525
Db 480 TNRNLGEFKLYPGGYLTCVPNSSSTGPQQLPDGVFVFASWVSRFYQLKPVGTAGPARGR 539
QY 526 LGRR 530
Db 540 LGRR 544

RESULT 6
Q9QTE7 PRELIMINARY; PRT; 544 AA.
AC Q9QTE7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Chiba virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=99565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Chiba 407;
RA Natori K., Suzuki K., Yamakawa Y., Tateumi M., Sakae K., Kobayashi S.,
RA Shinozaki K., Ishiko H., Miyamura T., Takeda N.;
RT "Expression and self-assembly of capsid proteins of the Chiba virus, a
RT genetically distinct Norwalk-like virus."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB028279; BAA82106.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58323 MW; ABDIC1FC4F93D872 CRC64;

Query Match 74.2%; Score 2091; DB 12; Length 544;
Best Local Similarity 72.74; Pred. No. 4.3e-162;
Matches 396; Conservative 53; Mismatches 80; Indels 16; Gaps 8;

QY 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMPDVAGSSAVATAGVNPIDPWIIINF 60
Db 1 MMASKDATPSADGATGAGQLVPEVNTADIPIDPVAGSSTALATAGVNLIDPWIIINF 60
QY 61 VQAPQGEFTISPNNTPGGVLDLSLGPLHNPFLHLSQMYNGVGNVRVIMLAGNAFTA 120
Db 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLPFLHLSQMYNGVGNVRVIMLAGNAFTA 120
QY 121 GKIIIVCIPGFGSHNLITIAQATLFPHVIAVRTLDPIEVLDPVLEDRVNLPHNNDRQOTM 180
Db 121 GKVIICCVPPGQSRISIAQATLFPHVIAVRTLDPIEVLDPVLEDRVNLPHNNDRQOTM 179
QY 181 RLVCMLYTPRLTGGGTG--DSFVAVGRVMTCPSPDFNLFVLPPTVEQKTRPFTLPL 238
Db 121 GKVIICCVPPGQSRISIAQATLFPHVIAVRTLDPIEVLDPVLEDRVNLPHNNDRQOTM 179
QY 239 SSLNSRAPLPIISGMGISPDNVQSVQFONGRCTLDGLRGVLTTPVLSLHVAKIRG--TSNG 296
Db 180 RLLCMLYTPRLTGGASGSDTSFVAVGRVLTCPGDFNLFVLPPTVEQKTRPFTVPIPL 239
QY 297 TVINLTDLGTPFPFEGPAPIGFPDGLGCDWHINMTQFGHSSQTDVDTT---PDT-- 351
Db 240 KYLSNRIPNPIEGMSLSPDQTNVQFNGRCTIDGQPLGTPVSVSQLCKFRGRTISGQ 299
QY 352 FVPHLGSII--QANGIGSGNYIGVLWSVSPSHPSGSQVDLWKIPNYGSSITEATHLAPSV 409

```

```

Db 360 FVPHLSSITLDENVSSGGDIYIGTQWTSPPSDGGANTNFWKIPDYGSSLAESQLAPAV 419
QY 410 YPGGGEVLVFFMSKIPGP--GAYSL-PCLLPQEIYISHLASEQAPTVEAALLHYVDPD 465
Db 420 YPGGNEVIIVFMASIPGPNQSGSNLVPCLLPQEIYTHFISEQAPIQGEAALLHYVDPD 479
QY 466 TGRTLGEFKAYPDGFLTCVPNGASSGPQQLPINGVVFVSWVSRFYQLKPVGTASSARGR 525
Db 480 TNRNLGEFKLYPGGYLTCVPNSSSTGPQQLPDGVFVFASWVSRFYQLKPVGTAGPARGR 539
QY 526 LGRR 530
Db 540 LGRR 544

RESULT 7
Q91185 PRELIMINARY; PRT; 544 AA.
AC Q91185;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Koblenz/433/2000/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=165252;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/Koblenz 433/2000/DE;
RA Kuenkel U., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated
RT with Norwalk-like viruses in Germany."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF394960; AAK72048.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58222 MW; 1B39A595B733A6A8 CRC64;

Query Match 73.9%; Score 2084; DB 12; Length 544;
Best Local Similarity 72.5%; Pred. No. 1.6e-161;
Matches 396; Conservative 51; Mismatches 81; Indels 18; Gaps 8;

QY 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMPDVAGSSAVATAGVNPIDPWIIINF 60
Db 1 MMASKDATPSADGATGAGQLVPEVNTAGPIPIDPVAGSSTALATAGVNLIDPWIIINF 60
QY 61 VQAPQGEFTISPNNTPGGVLDLSLGPLHNPFLHLSQMYNGVGNVRVIMLAGNAFTA 120
Db 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLPFLHLSQMYNGVGNVRVIMLAGNAFTA 120
QY 121 GKIIIVCIPGFGSHNLITIAQATLFPHVIAVRTLDPIEVLDPVLEDRVNLPHNNDRQOTM 180
Db 121 GKVIICCVPPGQSRISIAQATLFPHVIAVRTLDPIEVLDPVLEDRVNLPHNNDRQOTM 179
QY 181 RLVCMLYTPRLTGGGTG--DSFVAVGRVMTCPSPDFNLFVLPPTVEQKTRPFTLPL 238
Db 180 RLLCMLYTPRLTGGASGSDTSFVAVGRVLTCPGDFNLFVLPPTVEQKTRPFTVPIPL 239
QY 239 SSLNSRAPLPIISGMGISPDNVQSVQFONGRCTLDGLRGVLTTPVLSLHVAKIRG--TSNG 296
Db 240 KYLSNRIPNPIEGMSLSPDQTNVQFNGRCTIDGQPLGTPVSVSQLCKFRGRTISGQ 299
QY 297 TVINLTDLGTPFPFEGPAPIGFPDGLGCDWHINMTQFGHSSQTDVDTT---PDT-- 350
Db 300 KVLNLTDLGSPFMFAAPAPAGFPDGLGCDWHIEMSKIPNSS--TQSNPIVNVSVKPNQ 358
QY 351 TFVPHLGSII--QANGIGSGNYIGVLWSVSPSHPSGSQVDLWKIPNYGSSITEATHLAPSV 408
Db 359 QVPHLSSITLDENVSSGGDIYIGTQWTSPPSDGGANTNFWKIPDYGSSLAESQLAPAV 418
QY 409 VTPPGGEVLVFFMSKIPGP--GAYSL-PCLLPQEIYISHLASEQAPTVEAALLHYVDP 464

```



```
Db 121 GKVIICVPPGQSRSLTSLAQATLPHVIADVRTLDPEVEPLEDVRNVLXHNND-TQPTM 179
QY 181 RLVCMLYTPLRGCGGTG--DSFVAGRVMTCPSPDFNLFVLPPTVEQKTRPFTLPNPL 238
Db 180 RLVCMLYTPLRGCGGTGDSFVAGRVMTCPSPDFNLFVLPPTVEQKTRPFTLPNPL 239
QY 239 SSLNSRAPLPISGMGICSPDNVQSQVQFQNGRCTLGRLVGTTPVSLSHVAKIRG--TSNG 296
Db 240 KYLSNSRIPNPIEGMSLSLSPDQTNVQFQNGRCTLGRLVGTTPVSLSHVAKIRG--TSNG 299
QY 297 TVINLTGDTGTPHPPFEGPAPICGFDLGGCDWHINMTQFGHSSQTYD-----VDTPD 350
Db 300 KYLSNSRIPNPIEGMSLSLSPDQTNVQFQNGRCTLGRLVGTTPVSLSHVAKIRG--TSNG 358
QY 351 TFVPHLGSII--QANGIGSGNYIGVLSWVSPSPHSGSQVDLWKIPNYGSSITEATHLAPS 408
Db 359 QFVPHLPSITLDENVVSSGGDIYGTQWTSPPSDSGGANTFWKIPDYGSSLAESQLAPA 418
QY 409 VYPPGFEVLVFMFSKIPGP---GAYSL-PCLLPQEIYSHLASEQAPTVEGAALLHYVDP 464
Db 419 VYPPGFEVLVFMFSKIPGP---GAYSL-PCLLPQEIYSHLASEQAPTVEGAALLHYVDP 478
QY 465 DTGRTLGEFKAYPDGFLTCVPGNASSGPGQPLPINGVVFVSWVSRFYQLKPVGTASSARG 524
Db 479 DTNRNLGEFKLYPGGYLTCVPSNSTGPGQPLDGVFVFASVWSRFYQLKPVGTASSARG 538
QY 525 RLGLRR 530
Db 539 RLGVRR 544
```

RESULT 10

```
Q918D1 PRELIMINARY; PRT; 544 AA.
ID Q918D1
AC Q918D1
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/New Orleans/266/1993/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171834;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/266/1993/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/266/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/266/1993/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414402; AAL12959.1;
DR InterPro; IPR004005; Calici.coat.
DR InterPro; IPR008975; Viral Cap_coat.
DR Pfam; PF00915; Calici.coat.1.
SQ SEQUENCE 544 AA; 59354 MW; 50387E99852E801D CRC64;
```

Query Match 73.2%; Score 2063; DB 12; Length 544;
Best Local Similarity 72.0%; Pred. No. 8.3e-160;

```
Matches 393; Conservative 50; Mismatches 85; Indels 18; Gaps 8;
QY 1 MMASKDATSSVDCAGAGOLVPEVNASDPLAMDVPAGSSATATAGQVNPIDPWIINNF 60
Db 1 MMASKDATSSVDCAGAGOLVPEVNASDPLAMDVPAGSSATATAGQVNPIDPWIINNF 60
QY 61 VQAPQGEFTTSPNNTPGCVLFDLSGLPHLAPFLHLSQMYNGWGNVRIMLAGNAFTA 120
Db 61 VQAPQGEFTTSPNNTPGCVLFDLSGLPHLAPFLHLSQMYNGWGNVRIMLAGNAFTA 120
QY 121 GKIVISCIPIPGFSGHNTIAQAATLPHVIADVRTLDPEVEPLEDVRNVLXHNND-TQPTM 180
Db 121 GKIVISCIPIPGFSGHNTIAQAATLPHVIADVRTLDPEVEPLEDVRNVLXHNND-TQPTM 180
QY 181 RLVCMLYTPLRGCGGTG--DSFVAGRVMTCPSPDFNLFVLPPTVEQKTRPFTLPNPL 238
Db 180 RLVCMLYTPLRGCGGTG--DSFVAGRVMTCPSPDFNLFVLPPTVEQKTRPFTLPNPL 238
QY 239 SSLNSRAPLPISGMGICSPDNVQSQVQFQNGRCTLGRLVGTTPVSLSHVAKIRG--TSNG 296
Db 240 KYLSNSRIPNPIEGMSLSLSPDQTNVQFQNGRCTLGRLVGTTPVSLSHVAKIRG--TSNG 299
QY 297 TVINLTGDTGTPHPPFEGPAPICGFDLGGCDWHINMTQFGHSSQTYD-----VDTPD 350
Db 297 TVINLTGDTGTPHPPFEGPAPICGFDLGGCDWHINMTQFGHSSQTYD-----VDTPD 350
QY 300 RALNLTGDSGPFMAFAAPAGPFDLGGCDWHIEMSKIPNSS--TQNNPIVINSVKPNSQ 358
Db 300 RALNLTGDSGPFMAFAAPAGPFDLGGCDWHIEMSKIPNSS--TQNNPIVINSVKPNSQ 358
QY 351 TFVPHLGSII--QANGIGSGNYIGVLSWVSPSPHSGSQVDLWKIPNYGSSITEATHLAPS 408
Db 359 QFVPHLPSITLDENVVSSGGDIYGTQWTSPPSDSGGANTFWKIPDYGSSLAESQLAPA 418
QY 409 VYPPGFEVLVFMFSKIPGP---GAYSL-PCLLPQEIYSHLASEQAPTVEGAALLHYVDP 464
Db 419 VYPPGFEVLVFMFSKIPGP---GAYSL-PCLLPQEIYSHLASEQAPTVEGAALLHYVDP 478
QY 465 DTGRTLGEFKAYPDGFLTCVPGNASSGPGQPLPINGVVFVSWVSRFYQLKPVGTASSARG 524
Db 479 DTNRNLGEFKLYPGGYLTCVPSNSTGPGQPLDGVFVFASVWSRFYQLKPVGTASSARG 538
QY 525 RLGLRR 530
Db 539 RLGVRR 544
```

RESULT 11

```
Q9IV41 PRELIMINARY; PRT; 544 AA.
ID Q9IV41
AC Q9IV41
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Valetta/95/Malta.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=122920;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/Valetta/95/Malta;
RX MEDLINE=20404983; PubMed=10949950;
RA Green J., Valls J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2004)
DR EMBL; AJ277616; CAB89097.1;
DR InterPro; IPR004005; Calici.coat.
DR InterPro; IPR008975; Viral Cap_coat.
DR Pfam; PF00915; Calici.coat; 1.
SQ SEQUENCE 544 AA; 58415 MW; B6741846BD82B6D1 CRC64;
```

Query Match 72.6%; Score 2046; DB 12; Length 544;
Best Local Similarity 71.6%; Pred. No. 2e-158;
Matches 391; Conservative 53; Mismatches 84; Indels 16; Gaps 8;

```
QY 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSTAVATAGQVNPDPINN 60
DB 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSTAVATAGQVNPDPINN 60
QY 61 VQAPQGEFTISPNNTPGGVLFDSLGPHLNPFLHLSQMYNGWGNMRVIRIMAGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGGVLFDSLGPHLNPFLHLSQMYNGWGNMRVIRIMAGNAFTA 120
QY 121 GKIIIVSCIPPGSGSHNLTAQATLPPHVIADVRTLDPIEVPLEDVRNVLPHNDNRNQTM 180
DB 121 GKIIIVSCIPPGSGSHNLTAQATLPPHVIADVRTLDPIEVPLEDVRNVLPHNDNRNQTM 180
QY 181 RLVCMLYPLRTGGGTG--DSFVAVAGVMTCPSPDNFLFLVPPVEQKTRFTLPLNPL 238
DB 181 RLVCMLYPLRTGGGTG--DSFVAVAGVMTCPSPDNFLFLVPPVEQKTRFTLPLNPL 238
QY 239 SLSNSRAPLPIISGMSISPDNVQSFQNGRCITDGLRLVGTTPVSLSHVAKIRG--TSNG 296
DB 239 SLSNSRAPLPIISGMSISPDNVQSFQNGRCITDGLRLVGTTPVSLSHVAKIRG--TSNG 296
QY 297 TVINLTDLGTGTFPHFEGEPAPIGFDPDLGGCDWHINMTQFGHSSQTYD-----VDTPD 350
DB 297 TVINLTDLGTGTFPHFEGEPAPIGFDPDLGGCDWHINMTQFGHSSQTYD-----VDTPD 350
QY 351 TFPVHLGSI--QANGIGSGNYIGVLSWVSPPHSGQSDVLWKIPNYGSSITEATHLAPS 408
DB 351 TFPVHLGSI--QANGIGSGNYIGVLSWVSPPHSGQSDVLWKIPNYGSSITEATHLAPS 408
QY 359 QVPHLSSITLIDENVSNGDYICTQWTSPPSDSGGANTFWKIPDYGSSLAESQLAPA 418
DB 359 QVPHLSSITLIDENVSNGDYICTQWTSPPSDSGGANTFWKIPDYGSSLAESQLAPA 418
QY 409 VYPPGFGVLVFMFSKIPG---GAYSL-PCLLPOEYISHLASEOAPTVEAALLHYVDP 464
DB 409 VYPPGFGVLVFMFSKIPG---GAYSL-PCLLPOEYISHLASEOAPTVEAALLHYVDP 464
QY 419 VYPLVNEVIVFMACIRPGNCGSNLVPCLSPQEIYTHFISEQAPQGEAALLHYVDP 478
DB 419 VYPLVNEVIVFMACIRPGNCGSNLVPCLSPQEIYTHFISEQAPQGEAALLHYVDP 478
QY 465 DTGRTLGFGKAYPDGFLTCVNGASSGPQOLPINGVVFVSVWSRYQLKPVGTASSARG 524
DB 465 DTGRTLGFGKAYPDGFLTCVNGASSGPQOLPINGVVFVSVWSRYQLKPVGTASSARG 524
QY 479 DTNRNLGEFKLYPGGYLICVCPNNSSTGQOLPDGVFVFASVWSRYQLKPVGTAGPARG 538
DB 479 DTNRNLGEFKLYPGGYLICVCPNNSSTGQOLPDGVFVFASVWSRYQLKPVGTAGPARG 538
QY 525 RLGLRR 530
DB 525 RLGLRR 530
QY 539 RLGLRR 544
DB 539 RLGLRR 544
```

RESULT 12

```
QY9S14 PRELIMINARY; PRT; 540 AA.
AC QY9S14;
DT 01-MAY-1999 (TremBLrel. 10, Created)
DT 01-MAY-1999 (TremBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Norwalk virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=11983;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Norwalk-like;
RX MEDLINE=20254531; PubMed=10795514;
RA Schreier E., Doering F., Kuenkel U.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated
with small round structured viruses in Germany in 1997/98.";
RL Arch. Virol. 145:443-453(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Norwalk-like;
RA Schreier E., Doering F., Kuenkel U.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF093797; AAC64603.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 540 AA; 58018 MW; 647843487A654CBE CRC64;
```

Query Match

72.5%; Score 2044.5; DB 12; Length 540;

```
Best Local Similarity 71.2%; Pred. No. 2.7e-158;
Matches 386; Conservative 59; Mismatches 82; Indels 15; Gaps 7;
QY 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSTAVATAGQVNPDPINN 60
DB 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSTAVATAGQVNPDPINN 60
QY 61 VQAPQGEFTISPNNTPGGVLFDSLGPHLNPFLHLSQMYNGWGNMRVIRIMAGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGGVLFDSLGPHLNPFLHLSQMYNGWGNMRVIRIMAGNAFTA 120
QY 121 GKIIIVSCIPPGSGSHNLTAQATLPPHVIADVRTLDPIEVPLEDVRNVLPHNDNRNQTM 180
DB 121 GKIIIVSCIPPGSGSHNLTAQATLPPHVIADVRTLDPIEVPLEDVRNVLPHNDNRNQTM 180
QY 181 RLVCMLYPLRTGGGTG--DSFVAVAGVMTCPSPDNFLFLVPPVEQKTRFTLPLNPL 238
DB 181 RLVCMLYPLRTGGGTG--DSFVAVAGVMTCPSPDNFLFLVPPVEQKTRFTLPLNPL 238
QY 239 SLSNSRAPLPIISGMSISPDNVQSFQNGRCITDGLRLVGTTPVSLSHVAKIRG--TSNG 296
DB 239 SLSNSRAPLPIISGMSISPDNVQSFQNGRCITDGLRLVGTTPVSLSHVAKIRG--TSNG 296
QY 297 TVINLTDLGTGTFPHFEGEPAPIGFDPDLGGCDWHINMTQFGHSSQTYD--TTPTDTRVP 354
DB 297 TVINLTDLGTGTFPHFEGEPAPIGFDPDLGGCDWHINMTQFGHSSQTYD--TTPTDTRVP 354
QY 301 RVLNLSYVDGTFPVPELESFAPVGPDPDGGCDWHVGTFFEARDDQDPQSNVTFATNDSSFPV 360
DB 301 RVLNLSYVDGTFPVPELESFAPVGPDPDGGCDWHVGTFFEARDDQDPQSNVTFATNDSSFPV 360
QY 355 HLGSI---QANGIGSGNYIGVLSWVSPPHSGQSDVLWKIPNYGSSITEATHLAPSVP 411
DB 355 HLGSI---QANGIGSGNYIGVLSWVSPPHSGQSDVLWKIPNYGSSITEATHLAPSVP 411
QY 361 YLGSISPHNGDGFHSGDIIGSLDWISAP--DGSALDWSIPKYGSSLPDVTHLAPAVFP 418
DB 361 YLGSISPHNGDGFHSGDIIGSLDWISAP--DGSALDWSIPKYGSSLPDVTHLAPAVFP 418
QY 412 PFGSEVLVFMFSKIPGPG---AYSLPCLLPQYISHLASEOAPTVEAALLHYVDPDTGR 468
DB 412 PFGSEVLVFMFSKIPGPG---AYSLPCLLPQYISHLASEOAPTVEAALLHYVDPDTGR 468
QY 419 PFGSEVLVFMFSKIPGPG---AYSLPCLLPQYISHLASEOAPTVEAALLHYVDPDTGR 478
DB 419 PFGSEVLVFMFSKIPGPG---AYSLPCLLPQYISHLASEOAPTVEAALLHYVDPDTGR 478
QY 469 TLGEFKAYPDGFLTCVNGASSGPQOLPINGVVFVSVWSRYQLKPVGTASSARGRLGL 528
DB 469 TLGEFKAYPDGFLTCVNGASSGPQOLPINGVVFVSVWSRYQLKPVGTASSARGRLGL 528
QY 479 NLGEFKLYPDGFMTCVNPNSISGSPQTLFINGVVFVSVWSRYQLKPVGTASSARGRLGL 537
DB 479 NLGEFKLYPDGFMTCVNPNSISGSPQTLFINGVVFVSVWSRYQLKPVGTASSARGRLGL 537
QY 529 RR 530
DB 529 RR 530
QY 538 RR 539
DB 538 RR 539
RESULT 13
QY9V36 PRELIMINARY; PRT; 544 AA.
AC QY9V36;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Thistlehall/90/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=122925;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/Thistlehall/90/UK;
RX MEDLINE=20404983; PubMed=10949950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein: diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL; AJ277621; CAB89102.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58676 MW; 4D261DFA42621EDE CRC64;
```

Query Match

72.5%; Score 2044; DB 12; Length 544;

Best Local Similarity 71.6%; Pred. No. 3e-158;
Matches 391; Conservative 53; Mismatches 84; Indels 18; Gaps 8;

```
QY 1 MMASKDATTSSVDGASGAGQLVPEVNASDPLAMDPAVGSSTAVATAGQVNPIDPHIINF 60
Db 1 MMASKDATTSSVDGASGAGQLVPEVNASDPLAMDPAVGSSTAVATAGQVNPIDPHIINF 60
QY 61 VQAPQGEFTISPNNTPGGVLFDLISGLPHLNPFLHLSQMYNGWGNMRVRIMLAGNAFTA 120
Db 61 VQAPQGEFTISPNNTPGGVLFDLISGLPHLNPFLHLSQMYNGWGNMRVRIMLAGNAFTA 120
QY 121 GKIIIVSCIPPGGSHNLTAQATLPHVIADVRTLDPIEVPLEDVNRNVLPHNDNRNQOTM 180
Db 121 GKIIIVSCIPPGGSHNLTAQATLPHVIADVRTLDPIEVPLEDVNRNVLPHNDNRNQOTM 180
QY 181 RLVCMLYTPRTGGGTG--DSFVAGRVMTCPSPDENFLFVPPPTVEQKTRFTPLNPL 238
Db 181 RLVCMLYTPRTGGGTG--DSFVAGRVMTCPSPDENFLFVPPPTVEQKTRFTPLNPL 238
QY 239 SSLSNSRAPLPIISGMGISPDNVQVQFQNGRCTLDGRLVGTTPVSLSHVAKIRG--TSNG 296
Db 239 SSLSNSRAPLPIISGMGISPDNVQVQFQNGRCTLDGRLVGTTPVSLSHVAKIRG--TSNG 296
QY 297 TVINTLDELTPPHPEGPAPIGFDPDLGGCDWHINMTQFGHSSQTYD-----VDTPPD 350
Db 297 TVINTLDELTPPHPEGPAPIGFDPDLGGCDWHINMTQFGHSSQTYD-----VDTPPD 350
QY 351 TFVPHLGGSI--QANGIGSGNYIGVLSWVSPSPSPSGQVDLWKLIPNYGSSITEATHLAPS 408
Db 351 TFVPHLGGSI--QANGIGSGNYIGVLSWVSPSPSPSGQVDLWKLIPNYGSSITEATHLAPS 408
QY 359 QFVPHLSSTLDENVLSGGDIYIGNIQWTLVLLVILXEPKTNFWKIPDYGSSLAESQALAPA 418
Db 359 QFVPHLSSTLDENVLSGGDIYIGNIQWTLVLLVILXEPKTNFWKIPDYGSSLAESQALAPA 418
QY 409 VYPPGGEVLVFMFSKIPGP--GAYSL-PCLLPQEVYISHLASQAPTVEAALLHYVDP 464
Db 409 VYPPGGEVLVFMFSKIPGP--GAYSL-PCLLPQEVYISHLASQAPTVEAALLHYVDP 464
QY 419 VYPPGNEVLVYFMASIPGNQSGPNLVPCLLPQEVYITHFISEQAPIQGEAALLHYVDP 478
Db 419 VYPPGNEVLVYFMASIPGNQSGPNLVPCLLPQEVYITHFISEQAPIQGEAALLHYVDP 478
QY 465 DTGRTLGEKAYPDGFLTCVPGNASSGQQLPINGVVFVSWVSRFYQKPKVGTASSARG 524
Db 465 DTGRTLGEKAYPDGFLTCVPGNASSGQQLPINGVVFVSWVSRFYQKPKVGTASSARG 524
QY 479 DTNRNLGEFLKPGVLTCTVPNSSTGQQLPLDGVFVFWASWVSRFYQKPKVGTAGPARG 538
Db 479 DTNRNLGEFLKPGVLTCTVPNSSTGQQLPLDGVFVFWASWVSRFYQKPKVGTAGPARG 538
QY 525 RLGLRR 530
Db 525 RLGLRR 530
QY 539 RLGVRR 544
Db 539 RLGVRR 544
```

RESULT 14

```
ID Q8V7J5 PRELIMINARY; PRT; 531 AA.
AC Q8V7J5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein (Fragment).
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
RN [1]
RP SEQUENCE FROM N.A.
RA Yoda T., Terano Y., Suzuki Y., Yamazaki K., Oishi I., Kuzuguchi T.,
RA Kawamoto H., Utagawa E., Takino K., Shibata T.;
RA "Characterization of Norwalk virus G1 Specific Monoclonal Antibodies
RT Generated against Escherichia coli Expressed Capsid Protein and
RT Reactivity of Two G1 Specific Monoclonal Antibodies toward G1
RT Recombinant Fragments.", the EMBL/GenBank/DBJ databases.
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB063325; BAB3500.1; -.
DR InterPro; IPR004005; Calici_coat.
DR DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
FT NON_TER 531
FT TER 531
SQ SEQUENCE 531 AA; 56747 MW; 9688132FDDCFB88F CRC64;
```

Query Match 72.0%; Score 2031; DB 12; Length 531;
Best Local Similarity 72.2%; Pred. No. 3.3e-157;

```
Matches 385; Conservative 50; Mismatches 80; Indels 18; Gaps 8;
QY 8 ATSSVDGASGAGQLVPEVNASDPLAMDPAVGSSTAVATAGQVNPIDPHIINFVQAPQGE 67
Db 1 ATSSVDGATGAGQLVPEVNTADPIPDVAGSSTALATAGQVNLIDPHIINFVQAPQGE 60
QY 68 FTISPNNTPGGVLFDLISGLPHLNPFLHLSQMYNGWGNMRVRIMLAGNAFTAAGKIIVSC 127
Db 61 FTISPNNTPGGVLFDLISGLPHLNPFLHLSQMYNGWGNMRVRIMLAGNAFTAAGKIIVSC 120
QY 128 IPPGFGSHNLTAQATLPHVIADVRTLDPIEVPLEDVNRNVLPHNDNRNQOTMRLVCMY 187
Db 121 VPPGFSRTLSIAQATLPHVIADVRTLDPIEVPLEDVNRNVLPHNDNRNQOTMRLVCMY 179
QY 188 TPLRTGGGTG--DSFVAGRVMTCPSPDENFLFVPPPTVEQKTRFTPLNPLSSLSNSR 245
Db 180 TPLRTGASGTTDSFVAGRVMTCPSPDENFLFVPPPTVEQKTRFTPLNPLSSLSNSR 239
QY 246 APLPISGMGISPDNVQVQFQNGRCTLDGRLVGTTPVSLSHVAKIRG--TSNGTVINLITE 303
Db 240 IPNIEGMSLSPDQTVNQVQFQNGRCTLDGRLVGTTPVSLSHVAKIRG--TSNGTVINLITE 299
QY 304 LDGTPPHPEGPAPIGFDPDLGGCDWHINMTQFGHSSQTYD-----VDTPPDVPHLG 357
Db 300 LDGSPFWAFAAPAPAGFPDLGSCDWHIEMSKIPNST--TQNNPIVTVDSVKPNQSQ 358
QY 358 SI--QANGIGSGNYIGVLSWVSPSPSPSGQVDLWKLIPNYGSSITEATHLAPSYPGFG 415
Db 359 SITDDNVSSGDIYIGNIQWTLVLLVILXEPKTNFWKIPDYGSSLAESQALAPAVYPPGN 418
QY 416 EVLVFMFSKIPGP--GAYSL-PCLLPQEVYISHLASQAPTVEAALLHYVDPDTRTLG 471
Db 419 EVLVFMASIPGNQSGPNLVPCLLPQEVYITHFISEQAPIQGEAALLHYVDPDTRNLG 478
QY 472 EFKAYPDGFLTCVPGNASSGQQLPINGVVFVSWVSRFYQKPKVGTASSARG 524
Db 479 EFKLYPGGYLTCTVPNSSTGQQLPLDGVFVFWASWVSRFYQKPKVGTAGPARG 531
```

RESULT 15

```
ID Q8JW44 PRELIMINARY; PRT; 546 AA.
AC Q8JW44;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
RN [1]
RP SEQUENCE FROM N.A.
RA Katayama K., Takeda N., Natori K.;
RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=Funabashi258;
RA Katayama K., Takeda N.;
RA "Genetic and antigenic relationship among Norwalk-like viruses.";
RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB078335; BAC05516.1; -.
DR InterPro; IPR004005; Calici_coat.
DR DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR SEQUENCE 546 AA; 58834 MW; 1965F054E2C481E6 CRC64;
```

Query Match 72.0%; Score 2029.5; DB 12; Length 546;
Best Local Similarity 69.4%; Pred. No. 4.6e-157;
Matches 379; Conservative 68; Mismatches 82; Indels 17; Gaps 6;

Qy	1	MMMA	KDA	TSS	VDG	AS	GAG	Q	L	V	P	E	V	N	A	S	D	P	L	A	M	P	V	A	G	S	T	A	V	A	T	A	G	V	N	P	I	D	P	W	I	N	N	F	60														
Db	1	MMMA	S	K	D	A	P	Q	A	D	G	A	S	G	A	Q	L	V	P	E	N	T	A	D	T	F	L	M	E	B	E	V	A	G	T	T	A	V	A	T	A	G	V	N	I	D	P	W	I	V	N	F	60						
Qy	61	VQ	A	P	Q	G	E	F	T	I	S	P	N	N	T	P	G	S	L	I	S	G	P	H	N	L	P	L	L	H	L	S	O	M	Y	N	G	W	C	N	M	E	V	R	I	M	L	A	G	N	A	F	T	120					
Db	61	VQ	S	P	Q	G	E	F	T	I	S	P	N	N	T	P	G	D	I	L	F	D	Q	L	G	P	H	N	L	P	F	L	S	H	L	S	O	M	Y	N	G	W	C	N	M	E	V	R	I	L	L	A	G	N	A	F	S	T	120
Qy	121	GK	I	V	C	S	I	P	C	P	G	F	G	S	H	N	L	T	I	A	O	A	T	L	F	P	H	V	I	A	D	R	T	L	D	P	I	E	V	L	E	D	V	R	N	L	F	H	N	N	D	R	O	O	T	M	180		
Db	121	GK	I	V	C	C	P	P	G	T	S	S	L	T	I	A	O	A	T	L	F	P	H	V	I	A	D	R	T	L	E	P	I	E	M	P	L	E	D	V	R	N	L	F	H	N	N	D	R	O	O	T	M	179					
Qy	181	RL	V	C	M	L	T	P	L	T	G	G	T	G	--	D	S	E	V	A	G	R	W	M	T	C	P	S	D	F	N	F	L	V	P	P	T	V	E	O	K	T	R	P	F	L	P	N	L	P	L	238							
Db	180	RL	V	C	M	L	T	P	L	T	G	G	S	N	D	S	F	V	A	G	R	V	L	T	A	P	S	S	D	F	S	F	L	V	P	P	T	I	E	O	K	T	R	A	F	T	V	N	I	P	L	239							
Qy	239	S	S	L	N	S	R	A	P	L	P	I	S	G	M	I	S	P	D	N	V	Q	S	V	O	F	O	N	G	R	C	T	L	D	G	R	L	V	G	T	T	P	S	V	L	S	H	V	A	K	I	R	G	T	S	N	--	G	296
Db	240	Q	T	L	N	S	R	F	P	S	L	I	Q	M	I	S	P	D	A	S	V	Q	V	O	N	G	R	C	L	I	D	G	L	G	T	T	P	A	T	S	G	O	L	F	R	V	R	G	I	N	O	A	299						
Qy	297	T	V	I	N	L	T	E	L	D	G	T	F	H	F	F	E	G	P	A	P	I	G	P	D	L	G	C	D	H	I	N	M	T	Q	F	G	S	S	O	T	-----	Q	D	V	D	T	T	P	D	T	351							
Db	300	R	T	L	N	L	T	E	V	D	G	K	P	F	A	D	S	P	A	P	V	G	P	D	F	G	K	D	H	M	R	I	S	K	T	P	N	N	T	S	S	G	D	P	M	R	S	V	S	V	Q	T	N	V	Q	359			
Qy	352	F	V	P	H	L	G	S	I	Q	A	N	G	I	--	G	S	G	N	T	I	G	V	L	S	W	V	P	S	H	P	S	G	S	O	V	D	L	W	K	I	P	N	Y	G	S	S	T	E	A	T	H	L	A	P	S	408		
Db	360	F	V	P	H	L	G	S	T	Q	F	D	E	V	N	H	T	G	D	I	G	T	E	I	W	I	S	Q	P	S	T	P	G	T	D	I	N	L	W	E	I	P	D	Y	G	S	L	S	Q	A	N	L	A	P	419				

Search completed: June 1, 2004, 13:53:10
Job time : 31.8574 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:36:02 ; Search time 10.5352 Seconds

(without alignments)
4985.230 Million cell updates/sec

Title: US-09-926-799-3

Perfect score: 2895

Sequence: 1 MMASKDAPQADGASGAGQ.....QLKPVGTASTARSLGVRR 546

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2887	99.7	546	2 B37491	major capsid prote
2	2030.5	70.1	530	2 B37471	capsid protein - N
3	1149.5	39.7	542	2 S60616	capsid protein - h
4	1111	38.4	539	2 S40111	capsid protein - h
5	341.5	11.8	576	2 A53982	capsid protein - E
6	325	11.2	2344	1 RRWRH	genome polyprotein
7	321	11.1	2344	2 S64740	genome polyprotein
8	318	11.0	2344	2 S55399	genome polyprotein
9	299	10.3	703	1 C48562	coat protein - San
10	293.5	10.1	702	1 A48562	coat protein - San
11	284	9.8	668	1 VCVWFC	coat protein - fel
12	275.5	9.5	668	2 JQ2354	capsid protein - f
13	275.5	9.5	668	2 JQ2356	capsid protein - f
14	268.5	9.3	671	1 VCVWF9	coat protein - fel
15	267	9.2	668	1 VCVWFF	coat protein - fel
16	145.5	5.0	2194	1 GNNYE7	genome polyprotein
17	145	5.0	2206	2 S03822	genome polyprotein
18	143	4.9	2206	1 GNNY4P	genome polyprotein
19	140.5	4.9	2206	1 GNNY27	genome polyprotein
20	139	4.8	3473	1 A46112	genome polyprotein
21	139	4.8	3473	2 S27927	polyprotein - rice
22	131.5	4.5	2205	1 GNNY2W	genome polyprotein
23	128.5	4.4	2207	1 GNNY5P	genome polyprotein
24	127.5	4.4	3085	2 T00327	polyprotein - infe
25	123	4.2	2175	1 GNNYBE	genome polyprotein
26	122.5	4.2	2209	1 GNNY3P	genome polyprotein
27	122.5	4.2	13055	2 T16580	hypothetical prote
28	121.5	4.2	2209	1 GNNY2P	genome polyprotein
29	121	4.2	1742	2 S76110	hypothetical prote

30	119	4.1	2164	1 GNNY89	genome polyprotein
31	119	4.1	2185	1 GNNYSV	genome polyprotein
32	119	4.1	2185	1 GNNYSH	genome polyprotein
33	117.5	4.1	2207	1 GNNY1P	genome polyprotein
34	117.5	4.1	3972	2 S75251	hypothetical prote
35	117	4.0	2207	2 S09553	genome polyprotein
36	116.5	4.0	2150	1 GNNYH2	genome polyprotein
37	116	4.0	1200	2 T17404	hyalin - sea urchi
38	116	4.0	2185	1 JQ2021	genome polyprotein
39	115	4.0	888	1 GNLJHD	pol polyprotein -
40	113.5	3.9	2333	1 GNNY2F	genome polyprotein
41	113	3.9	896	2 I56563	interleukin-3 rece
42	112.5	3.9	708	1 M2XRLL	structural protein
43	112.5	3.9	2214	1 A48548	genome polyprotein
44	112	3.9	2179	1 GNNYH4	genome polyprotein
45	111.5	3.9	1344	2 T14316	rig-1 protein - mo

ALIGNMENTS

RESULT 1

B37491 major capsid protein [similarity] - Southampton virus

N:Alternate names: orf2 protein

C:Species: Southampton virus

C:Date: 22-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 18-Aug-2000

C:Accession: B37491

RefSeq: P. R. J. C. Ashley, C. R. Clarke, I. N. J.

Sequence 259,546,519,520,521,522,523

Article: Sequence and genome organization of a human small round-structured (Norwalk-like

A:Reference number: A37491; PMID:93142023; PMID:8380940

A:Accession: B37491

A:Status: not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-546 <LM>

A:Cross-references: GB:L07418; NID:gl236787; PIDN:AAA92984.1; PID:g295114

A:Note: sequence extracted from NCBI backbone (NCBI:123458)

A:Note: small round-structured virus, SRSV, Norwalk virus, Norwalk-like virus, serotype

C:Superfamily: human calicivirus capsid protein

C:Keywords: glycoprotein

F:303,340,441/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 99.7%; Score 2887; DB 2; Length 546;

Best Local Similarity 99.6%; Pred. No. 2.9e-207;

Matches 544; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	MMASKDAPQADGASGAGQLVPEVNTADPLPMEPVAGTTAVATAGQVNMIDPWVNNF	60
DB	1	MMASKDAPQADGASGAGQLVPEVNTADPLPMEPVAGTTAVATAGQVNMIDPWVNNF	60
QY	61	VQSPQGETTIGPNTTGGDILFDLQGLPHLNPFLSHLSQVNGVGNMVRILLAGNAFSA	120
DB	61	VQSPQGETTIGPNTTGGDILFDLQGLPHLNPFLSHLSQVNGVGNMVRILLAGNAFSA	120
QY	121	GKIIIVCCVPPGTSSTSLTIAQATLPPHVIADVTLEPIEMPLEDVRNVLHYHTNDNQPMR	180
DB	121	GKIIIVCCVPPGTSSTSLTIAQATLPPHVIADVTLEPIEMPLEDVRNVLHYHTNDNQPMR	180
QY	181	LVCMLYTLRTGGGSGNSDSFVAGRVLTAPSSDSFLFLVPTTIEQKTRAFVNPILQ	240
DB	181	LVCMLYTLRTGGGSGNSDSFVAGRVLTAPSSDSFLFLVPTTIEQKTRAFVNPILQ	240
QY	241	TLNSRRFSLTQGMILSPDASQVQFQNGRCLIDGQLLGTTPATSGQLFVRGKKNQGAR	300
DB	241	TLNSRRFSLTQGMILSPDASQVQFQNGRCLIDGQLLGTTPATSGQLFVRGKKNQGAR	300
QY	301	TLNLTVEVDGKPFMAFDSAPVGFDPFGKCDWHMRISKTPNNNTSSGDPMRSVSVQTNVQGF	360
DB	301	TLNLTVEVDGKPFMAFDSAPVGFDPFGKCDWHMRISKTPNNNTSSGDPMRSVSVQTNVQGF	360
QY	361	VPHLGSIQFDEVFNHTDGTGTIWIISQPSPTPGTDTNLWEIPDYGSSLSQANLAPPV	420

```

361 VPHLSIQFDEVFNHPTGDYIGTIEWISQSTPTEGTDINLWEIPDYGSSLSQAANLAPPV 420
421 PPGGEGEALVYFVSAPFPNNRSANDVPCLLPOEYITHFVSEQAPTMGDAALHYVDPD 480
421 PPGGEGEALVYFVSAPFPNNRSANDVPCLLPOEYITHFVSEQAPTMGDAALHYVDPD 480
481 TNRNLGEFKLYPGGYLTICVPGVGAGPQQLPLNGVFLFVSVWSRFPYQLKPVGTASTARS 540
481 TNRNLGEFKLYPGGYLTICVPGVGAGPQQLPLNGVFLFVSVWSRFPYQLKPVGTASTARG 540
541 LGVRR 546
541 LGVRR 546

RESULT 2
B37471
capsid protein - Norwalk virus
C:Species: Norwalk virus
C:Date: 24-Feb-1994 #sequence_revision 25-Apr-1997 #text_change 28-Jul-2000
C:Accession: B37471
R:Jiang, X.; Wang, M.; Wang, K.; Estee, M.K.
Virology 195, 51-61, 1993
A:Title: Sequence and genomic organization of Norwalk virus.
A:Reference number: A37471; MUID:93303939; PMID:8391187
A:Accession: B37471
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: genomic RNA
A:Residues: 1-530 <JIA>
A:Cross-references: GB:M87661, NID:G1061311; PIDN:AAB50466.1; PID:G1061313
A:Note: sequence extracted from NCBI backbone (NCBI:P134157)
C:Superfamily: human calicivirus capsid protein

Query Match 70.1%; Score 2030.5; DB 2; Length 530;
Best Local Similarity 69.2%; Pred. No. 1.8e-143;
Matches 378; Conservative 70; Mismatches 81; Indels 17; Gaps 6;

QY 1 MMASKADAPQASAGAGAGOLVPEVNTADPLMEPVAGPTTAVATAGQVNMIDPMTVNNF 60
DB 1 MMASKADATSSVDCASGAGOLVPEVNTADPLMEPVAGPTTAVATAGQVNMIDPMTVNNF 60
61 VQSPQGEFTTSPNNTPGDILFDLQGLPHLPFLSHLSQMYNGVGNMVRILLAGNAFSA 120
61 VQAPQGEFTTSPNNTPGDILFDLQGLPHLPFLSHLSQMYNGVGNMVRILLAGNAFSA 120
121 GKTIIVCCVPPGFTSSSLTIAQATLFPHVIAVDTLEPIEMPLEDVRNLVHTND-NQPTM 179
121 GKTIIVCCVPPGFTSSSLTIAQATLFPHVIAVDTLEPIEMPLEDVRNLVHTND-NQPTM 180
180 RLVCMLYTPLRTGGSGNSDSFVAGVLTAPSSDFSLFLVPTTIEQKTRAFVTPNIPL 239
181 RLVCMLYTPLRTGGSGTG--DSFVAGVMTCPSPDFNPLFLVPTTIEQKTRPFTL 238
240 QTLNSRPSLIQGMILSPDASQVVFQNGRCLLDGQLGTTTATSPATSOQLFRVKGKINQA 299
239 SSLNSRAPLPIISIGISPDNVQSVQFQNGRCLLDGRLVGTTPVSLSHVAKIRGTSN--G 296
300 RTLNLTEVDGPKFMAFSPAPVGPDPFGKCDWHMIRISKTPNNTSSGDPMSVSQVTVNQ 359
297 TVINLTEDGTPFPFEPGAPIGPFDLGGCDWHINMTQFGHSSQT-----QYDVDTTPTD 351
360 FVPHLGSTOFDEVFNHPTGDYIGTIEWISQSTPTEGTDINLWEIPDYGSSLSQAANLAPP 419
352 FVPHLGSIQANGI-----GSGNVGVLSWISPSHSPSGSQVDLWKLPNYGSSITTEATHLAPS 408
420 VPPGFGGALVYFVSAPFPNNRSAPNDVPCLLPOEYITHFVSEQAPTMGDAALHYVD 479
409 VYPPGFGGALVYFVSMKMPG-----GAYNLPCLLPQEIYIHLASEQAPTVGEAALHYVD 464
480 DTNRNLGEFKLYPGGYLTICVPGVGAGPQQLPLNGVFLFVSVWSRFPYQLKPVGTASTARS 539
465 DTGRNLGEFKAYPDGDLTICVPGVGAGSGPQQLPLNGVFLFVSVWSRFPYQLKPVGTASSARG 524

```

```

QY 540 RLGVRR 545
DB 525 RLGLRR 530

RESULT 3
S60616
capsid protein - human calicivirus (strain Melksham)
C:Species: human calicivirus
A:Variety: strain Melksham
C:Date: 23-May-1997 #sequence_revision 23-May-1997 #text_change 28-Jul-2000
C:Accession: S60616
R:Green, S.M.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
Virus Res 37, 271-283, 1995
A:Title: Capsid diversity in small round-structured viruses: molecular characterization
A:Reference number: S60615; MUID:96136658; PMID:8533462
A:Accession: S60616
A:Molecule type: genomic RNA
A:Residues: 1-542 <GRE>
A:Cross-references: EMBL:X81879; NID:G976077; PIDN:CAA57462.1; PID:G976079
A:Experimental source: strain Melksham
A:Note: it is uncertain whether Met-1 or Met-3 is the initiator
C:Superfamily: human calicivirus capsid protein
C:Keywords: capsid protein; coat protein

Query Match 39.7%; Score 1149.5; DB 2; Length 542;
Best Local Similarity 44.7%; Pred. No. 8.4e-78;
Matches 256; Conservative 82; Mismatches 176; Indels 59; Gaps 19;

QY 1 MMASKADAPQASAGAGAGOLVPEVNTADPLMEPVAGPTTAVATAGQVNMIDPMTVNNF 60
DB 1 MMASKADAPSTDGAAG---LVPESNN-EVMALPEVAGAAALAAPVTGQTNIIDPWIRANF 56
61 VQSPQGEFTTSPNNTPGDILFDLQGLPHLPFLSHLSQMYNGVGNMVRILLAGNAFSA 120
61 VQAPNGEFTVSPNAPQEVLLNLELPELNPYLALHARMYNGVAGMEVQVMLAGNAFTA 116
121 GKTIIVCCVPPGFTSSSLTIAQATLFPHVIAVDTLEPIEMPLEDVRNLVHTND-NQPTM 179
117 GKLVFAAVPPHPFVENLSPQKITMFPFHVIIIDVLTLEPVLPLPDVRNSFFHYNQKDDPKM 176
180 RLVCMLYTPLRTGGSGNSDSFVAGVLTAPSSDFSLFLVPTTIEQKTRAFVTPNIPL 239
177 RIVANLYTPLRS-NGSGD-DVFTVSCRVLTRPSPDFDTYLVPTTVESKTKPFTLPILT 234
240 QTLNSRPSLIQGMILSPDASQVVFQNGRCLLDGQLGTTTATSPATSOQLFRVKGKI----- 295
235 GELNSRFPVPIQMYTSPNEVISVQCNGRCLTDGELQGTTLQVSGICAFKGEVTAHL 294
296 --NQGARTLNLTEDVGKPF-MAFSPAPVGPDPF-GKCDWHMIRISKTPNNTSSGDPM--- 348
295 HDNDHLNVTITNLNGSPFDESDIPAPLGVDPFQGRVGVISORDKQNAAGHSEPANRG 354
349 RSVSVQTVNQGVFVPHLGSIQFDEVFNHPTGDYIGTIEW-----ISQP--STPPGTD--- 397
355 HDVAVPTTYAQYTPKLGQIQ-----IGT--WQTDLDLTWNQVKEFTPVGLNDTE 400
398 -INLWEIPDYGSSLSQAANLAP---VPPPGFGEALVYFVSAPFPNNRSAPNDVPCLLP 453
401 HFNQWVPVRYAGALNLTNLAAPSAPVFPF---GEFLFERSHLPLKGGYGNP-AIDCLLP 456
454 QEYIYTHFVSEQAPTMGDAALHYVDPTNRNLGEFKLYPGGYLTICVPGVGAGPQQLPLN 513
457 QEWVQHFYQEAAPSSEVALVRIINFDTCRALFEAKLHRAGMTVSSN--TSAPVVVPAN 514
514 GVFLFVSVWSRFPYQLKPVGTASTARSRLGVRR 546
515 GYFRFDSWNQFYSLAPMGTGN-----GRRRV 541

RESULT 4
S40111
capsid protein - human calicivirus (strain Bristol isolate B493)

```

C;Species: human calicivirus
A;Variety: strain Bristol isolate B493
C;Date: 25-Dec-1994 #sequence_revision 27-Feb-1997 #text_change 28-Jul-2000
R;Green, S.M.; Dingle, K.E.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
submitted to the EMBL Data Library, December 1993
A;Description: Human enteric caliciviridae: a new prevalent SRSV group defined by RNA-de
A;Reference number: S40111
A;Accession: S40111
A;Molecule type: genomic RNA
A;Residues: 1-539 <GRE>
A;Cross-references: NID:X76716; NID:g436410; PIDN:CRA54134.1; PID:g436411
A;Experimental source: human enteric calicivirus strain Bristol isolate B493
C;Superfamily: human calicivirus capsid protein
C;Keywords: capsid protein; coat protein

Query Match 38.4%; Score 1111; DB 2; Length 539;
Best Local Similarity 44.3%; Pred.No. 6.1e-75;
Matches 250; Conservative 76; Mismatches 190; Indels 48; Gaps 17;

Qy 1 MMWASKDAPQSADGASGAGOLVPEVNATDPLMEPVAGTAVATAGQNMDIPWVNNF 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MKMASNDANPS-DGS--AAALVFEVNV-EVMALPEVVGAIAAPVAGQQNVDPWRNF 56
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 61 VQSPQGEFTISPNNTPGDILFDIQLGPHNLPIFLSHLSOMYNGVGNRVIRILLAGNAFS 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 57 VOAPGGEFTVSPRNAPGEILWSAPGLDNLPYLHLSRMVNYGAYGGFEVQVILAGNAFTA 116
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 121 GKIIICVPGETSSSITTAQTLFGHVHTADVRLTEPIEMPLEDVENVLYHTND-NQPTM 179
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 117 GKVFIAAPVNPFFTEGLSPSQVTMFPHIIVDRQLSEVLIPDPVRNNFYHNQANDSTL 176
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 180 RLCLMYLTPLRTCGSGNSDSFVACRGVLTAPSDSFSLFPVPTIEQKTRAFTPNIPL 239
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 177 KLIALMYTPLR--ANNAGDDVFTVSCRVLTRSPDFDFILVPPTVESRKFTVPVLTV 234
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 240 QTLNSGRFSLIOGMILSPDASOVQFQNGRCILIDGOLLGTTPATSGQLFRVGKINOGA 299
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 235 EEMSNGRFPIPLEKLYTGSPSAFVQPQNGRCTTDGVLGTTQLSAVINCFNRGDVTHIA 294
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 300 ----RTLNLTEVDGKFPMAFDS-PAPVGPDF-GKCDMHMRISKTNNTSSGD---PMRS 350
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 295 GSHDYTMNLASQWNSYDTEELPAPLGTFDFVGK-----IQGLLTQTTRADGSTRHK 348
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 351 VSQTVNQGVFPHLGSIGQFDEVFNHPTGYD-----IGTIEMISQSPTPPGTDINL 400
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 349 ATYSTGSVHFTPKLGSVQFTDTNN---DFQAQNTKFTPVGVIQGDHHQNEP----QQ 401
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 401 WEIPDYSSLSQAANLAPPVPPGFGEALVYVSFAFPGNRRSAPN-DVPCLLLPQSYITH 459
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 402 WLLPNYSGRTHGNVHLAPAAPTFFGEQLLFRRSTMPGCS--GYPNMNLDCLLPQEWVLH 459
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 460 FYSEQAPTWGDAALLHYVDPDTRNIGEFKLPGGYLTCVPNGVGAGPOO--LPLNGVPL 517
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 460 FYQEAPPAOSDVALFRVNPDTGRVLFECKLHKSQYITV----AHTGPYDLVLPNGPYR 515
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 518 FVSWSRFYOLKPVGVTASTARSRL 541
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 516 FDSWNVQFTYLAPEGNGTGRRAL 539
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 5
A53982
capsid protein - European brown hare syndrome virus
C;Species: European brown hare syndrome virus
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 28-Jul-2000
R;Wirbach, C.; Meyers, G.; Ohlinger, V.F.; Capucci, L.; Eskens, U.; Haas, B.; Thiel, H.
J. Virol. 68, 5164-5173, 1994
A;Title: European brown hare syndrome virus: relationship to rabbit hemorrhagic disease
A;Reference number: A53982; MUID:94309183; PMID:7518531
A;Accession: A53982
A;Status: preliminary

A:Molecule type: mRNA
A:Residues: 1-576 <WIR>
A:Cross-references: GB:U09199
C:Superfamily: human calicivirus capsid protein

Query Match 11.8%; Score 341.5; DB 2; Length 576;
Best Local Similarity 25.1%; Pred. No. 1.5e-17;
Matches 142; Conservative 66; Mismatches 217; Indels 141; Gaps 23;

QY 7 DAQSDAGSAGAGQLV-----PEVNTADPLPMEPVAGPTTAVATAG-----QVN 50
DB 8 DAGTATTASVPGTTDGMDPGVASDVTADNV-----AASVATAGIGGPPQAS 59
QY 51 MIDPWVNNFVQSQGEFTISPNNTPGDILFDLQLGPHLNPFSLHSQMYNGWGNMKRV 110
DB 60 PQSSWRVNFYND---VFTSVTDAPGSILYSVQHSQNNPFTQVLQSVYAGWAGGQMR 116
QY 111 ILLAGNAPSAGKIIVCVPPGF-TSSSLTTAQTALFPHVIADVRLTPIEMPLEDVRLV 169
DB 117 FIVAGSGIFGRLVCAIIPGIIQIQGLEVRQ---FPHVIDARSLSEPVITMPDLRPEM 173
QY 170 YH-TNDNQPMRLVCMLYTLRLTRGGSGNSDFVAGRVLTAPSSDFSFLFLVPPTIEQ 228
DB 174 YHFTGDPGLVPTLVSVYNNLINPFGGTTG---AIQVTVETRSEDFEVLIRAPS--SK 228
QY 229 TRAFVTVPNIPLQTL-----SNSRFPSLIQGMILSPDASQVQFQNGRCCLIDGQLLGTTP 282
DB 229 TVDSVNPSSMLLTPVLVTGAGSDNRWGAPIVGLQVPVGGFST---SNRHNWNGSTYGWSS 285
QY 283 ATSGQLFRVRGKI-----NOGARTLNLTEVDGKPFNAFDSAPVGPDPDF 326
DB 286 PRPDDIDHPGSGNYSYPTGSATNTIETWYANAGTATTN-----PISNIAPDGPDDM 335
QY 327 GKCDWHRISKTPNNNTSSGDPMSVSQTVNVQGVPHLGSIQPDEVFNHPTG-DYIGTIE 385
DB 336 GAI-----PFGSTTIPTGAW-----VGFGQVWNASNGTPIYGVTVQ 370
QY 386 -----WISQFSTPGPTDINLWEIPDYGSSLSQAANLAPPVF-----PFGFG 426
DB 371 AVELGFANGAPPSIRPVTITTTGAQLVAKSI--YGVATQNOSSAGIIFLSKGMVSTPG-- 426
QY 427 EALVYFVSAPPGNNRSAPNDVCLLIPQEVITHFVSEQAPTMGDAAALLHYVDPTNRNLG 486
DB 427 ---VAATTTTPQSAIVTTPGTVAAP-----IGKNTPIMFSAVV-----RRTG 467
QY 487 EPKLYPGGYLTCVPGNVGAGPQQLPL 512
DB 468 DVNAGPES-VNGTQYGVGSQPLSVTL 492

RESULT 6
RRWRH

genome polyprotein - rabbit hemorrhagic disease virus
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C:Species: rabbit hemorrhagic disease virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_revision 29-May-1998
C:Accession: A41039
R:Meysers, G.; Wirblich, C.; Thiel, H.J.
Virology 184, 663-676, 1991
A:Title: Rabbit hemorrhagic disease virus-molecular cloning and nucleotide sequence
A:Reference number: A41039; PMID:91361557; PMID:1840711
A:Accession: A41039
A:Molecule type: genomic RNA
A:Residues: 1-2344 <MEY>
A:Cross-references: GB:M67473
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: nucleotidyltransferase

Query Match 11.2%; Score 325; DB 1; Length 2344;
Best Local Similarity 26.7%; Pred. No. 1.9e-15;
Matches 118; Conservative 57; Mismatches 183; Indels 84; Gaps 18;

QY 5 SKDAPQSDAGSAGAGQLVPEVNTADPLPMEPVAGPTTAV-----ATAG-----Q 48

Db 1770 ARAAPQGEAAGTATTASVPGTTTGDG---MDPGVAVTTSVITAENSASIASIATAGIGGPPQQ 1826
QY 49 VNMIDPWI VNNFVQS PQSGEFTTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWVGNMR 108
Db 1827 VDOQETWRNFYND---VFTWSVADAPGSILYTVQHSFQNNPFTAVLSQMYAGWAGMQ 1883
QY 109 VRILLAGNAFSAKIIIVCCVPPGF-TSSSLTTIAQATLPHVIAVRTLEPIEMPLEDVNR 167
Db 1884 FRFIVAGSVFGRLVRAVIPPGEIIGPGLVRQ---FPHVVIDARSLEPVTITMPDLRP 1940
QY 168 VLYH-TNDNQPTMRLVCMLYTLRTGGSGNSDSFVAVGRVLTAPSSDFSLFLVPP--- 223
Db 1941 NMYHPTGDPCLVPTLVLSVNNLINFPGGSTS---AIQVTVETRSEDFEFVIRAPSSK 1997
QY 224 TIEQKTRAFVTPNIPQLTSLN-SRFPSLIQGMILSPDASQVVFQNGRCLIDGQLLGTTP 282
Db 1998 TVDSISPAGLLTTPVLTVGVNDNRWNGQIVGLQVPVGGFSTC---NRHWNLNGSTVGWSS 2054
QY 283 ATSGQLFRVRGKIN-QGARTLNLTETV-DGKPFMAFDSP-----APVGFDPFGKCDWHMRIS 336
Db 2055 PRFGDIDHRRGSASVSGSNATNLQFWYANAGSAIDNPISQVAPDGFDP-----MS 2105
QY 337 KTPNNTSSGDPMSVSQVTVNQGVFPHLGSIQFDEVNHPGTGDIYGTIEWISQSPSTPGT 396
Db 2106 FVPFNGPG-----IPAGWVGFGAI-----WNSNGAPNVT 2136
QY 397 DINLWEIPDYGSSLSQAANLAP 418
Db 2137 TVQAYEL---GFATGAPGNLQ 2155
RESULT 7
S64740
genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)
N;Contains: VP60 protein
C;Species: rabbit hemorrhagic disease virus
A;Variety: isolate AST/89
C;Date: 12-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Jun-2000
C;Accession: S64740; S46944; S49018; S65012
R;Casais, R.; Martin-Alonso, J.; Boga, J.; Parra, F.
submitted to the EMBL Data Library, May 1995
A;Description: Genomic organization of rabbit hemorrhagic disease virus determined by di
A;Reference number: S64740
A;Accession: S64740
A;Molecule type: genomic RNA
A;Residues: 1-2344 <CAS>
A;Cross-references: EMBL:Z49271; NID:g1182032; PIDN:CAA89265.1; PID:g1150552
A;Experimental source: isolate AST/89
R;Boga, J.; Casais, R.; Marin, M.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra, R
submitted to the EMBL Data Library, July 1993
A;Description: Molecular cloning, sequence and expression of the capsid protein gene fr
A;Reference number: S46944
A;Accession: S46944
A;Molecule type: genomic RNA
A;Residues: 1650-2344 <BOC>
A;Cross-references: EMBL:Z24757; NID:g515622; PIDN:CAA80881.1; PID:g515623
A;Experimental source: isolate AST/89
R;Parra, F.; Boga, J.A.; Marin, M.S.; Casais, R.
Virus Res. 27, 219-228, 1993
A;Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus supp
A;Reference number: S49018; MUID:93255896; PMID:8488721
A;Accession: S49018
A;Molecule type: genomic RNA
A;Residues: 1650-1796 <PAR>
A;Cross-references: EMBL:Z24757
A;Experimental source: isolate AST/89
A;Accession: S65012
A;Molecule type: protein
A;Residues: 1767-1793;1875-1877, 'X', 1879-1881;1936-1938, 'X', 1940-1941 <PAW>
C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C;Keywords: blocked amino end; polyprotein

Query Match 11.1%; Score 321; DB 2; Length 2344;
Best Local Similarity 26.7%; Pred. No. 3.8e-15;
Matches 116; Conservative 55; Mismatches 185; Indels 84; Gaps 18;
QY 5 SKDAPQADGASGAGQLVPEVNTADPLPMEPVAGPTTAV-----ATAG-----Q 48
Db 1770 ARTAPQGEAAGTATTASVPGTTTGDG---MDPGVAVTTSVITAENSASIASIATAGIGGPPQQ 1826
QY 49 VNMIDPWI VNNFVQS PQSGEFTTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWVGNMR 108
Db 1827 VDOQETWRNFYND---VFTWSVADAPGSILYTVQHSFQNNPFTAVLSQMYAGWAGMQ 1883
QY 109 VRILLAGNAFSAKIIIVCCVPPGF-TSSSLTTIAQATLPHVIAVRTLEPIEMPLEDVNR 167
Db 1884 FRFIVAGSVFGRLVRAVIPPGEIIGPGLVRQ---FPHVVIDARSLEPVTITMPDLRP 1940
QY 168 VLYH-TNDNQPTMRLVCMLYTLRTGGSGNSDSFVAVGRVLTAPSSDFSLFLVPP--- 223
Db 1941 NMYHPTGDPCLVPTLVLSVNNLINFPGGSTS---AIQVTVETRSEDFEFVIRAPSSK 1997
QY 224 TIEQKTRAFVTPNIPQLTSLN-SRFPSLIQGMILSPDASQVVFQNGRCLIDGQLLGTTP 282
Db 1998 TVDSISPAGLLTTPVLTVGVNDNRWNGQIVGLQVPVGGFSTC---NRHWNLNGSTVGWSS 2054
QY 283 ATSGQLFRVRGKIN-QGARTLNLTETV-DGKPFMAFDSP-----APVGFDPFGKCDWHMRIS 336
Db 2055 PRFGDIDHRRGSASVSGSNATNLQFWYANAGSAIDNPISQVAPDGFDP-----MS 2105
QY 337 KTPNNTSSGDPMSVSQVTVNQGVFPHLGSIQFDEVNHPGTGDIYGTIEWISQSPSTPGT 396
Db 2106 FVPFNGPG-----IPAGWVGFGAI-----WNSNGAPNVT 2136
QY 397 DINLWEIPDYGSSLSQAANLAP 418
Db 2137 TVQAYEL---GFATGAPGNLQ 2155
RESULT 8
S55399
genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)
C;Species: rabbit hemorrhagic disease virus
A;Variety: isolate BS89
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 18-Jun-1999
C;Accession: S55399
R;Rossi, C.
submitted to the EMBL Data Library, May 1995
A;Reference number: S55399
A;Accession: S55399
A;Molecule type: genomic RNA
A;Residues: 1-2344 <ROS>
A;Cross-references: EMBL:X87607; NID:g854640; PIDN:CAA60910.1; PID:g854641
A;Experimental source: isolate BS89
C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C;Keywords: polyprotein
Query Match 11.0%; Score 318; DB 2; Length 2344;
Best Local Similarity 26.2%; Pred. No. 6.4e-15;
Matches 116; Conservative 58; Mismatches 184; Indels 84; Gaps 18;
QY 5 SKDAPQADGASGAGQLVPEVNTADPLPMEPVAGPTTAV-----ATAG-----Q 48
Db 1770 ARTAPQGEAAGTATTASVPGTTTGDG---LDPGVAVTTSVITAENSASIASIATAGIGGPPQQ 1826
QY 49 VNMIDPWI VNNFVQS PQSGEFTTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWVGNMR 108
Db 1827 VDOQETWRNFYND---VFTWSVADAPGSILYTVQHSFQNNPFTAVLSQMYAGWAGMQ 1883
QY 109 VRILLAGNAFSAKIIIVCCVPPGF-TSSSLTTIAQATLPHVIAVRTLEPIEMPLEDVNR 167
Db 1884 FRFIVAGSVFGRLVRAVIPPGEIIGPGLVRQ---FPHVVIDARSLEPVTITMPDLRP 1940
QY 168 VLYH-TNDNQPTMRLVCMLYTLRTGGSGNSDSFVAVGRVLTAPSSDFSLFLVPP--- 223

```
Db 1941 NMYHPTGDPGLVTLVSVNNLINPFGGSTS---AIQVTVETRPSEDFEFVIRAPSSK 1997
QY 224 TIEQKTRAPTVNPIQLTSLN-SRPSLLIQMILSPDASQVVOFQNGRLIDQLLGTTP 282
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1998 TVDSISPAGLLTPTVLTVGNDNRWNGQIVGLQVPFGGSTC---NRHNLNGSTVGWSS 2054
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 283 ATSGQLFRVRGKIN-QGARTLNTEV-DGKPFMAEDSP-----APVGFDPFGKCDHMRIS 336
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2055 PRFADIDHRGSASYPGSNATVLOFWYANAGSVDNPIISQVAPDGFDP-----MS 2105
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 337 KTPNNNTSSGDMRMSVSVQTVNQVFPHLGSIQFDEVFNHPTGDYGTIEWISQSPTPPCT 396
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2106 FYPFNGPG-----TPAGWVGFAL-----WNSNGAPNVT 2136
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 397 DINLWEIPDYGSSLSQANLAP 418
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2137 TVQAVEL---GPATGAPGNLQ 2155
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
C48562
coat protein - San Miguel sea lion virus (serotype 4)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: C48562
R:Neill, J.D.
Virus Res. 24, 211-222, 1992
A:Title: Nucleotide sequence of the capsid protein gene of San Miguel sea
eins.
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Accession: C48562
A:Molecule type: genomic RNA
A:Residues: 1-703 <NEI>
A:Cross-references: GB:M87482; NID:g334886; PIDN:AAAL6220.1; PID:g334888
A:Note: sequence extracted from NCBI backbone (NCBIP:113567)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.3%; Score 299; DB 1; Length 703;
Best Local Similarity 27.2%; Pred. No. 2.9e-14;
Matches 108; Conservative 60; Mismatches 173; Indels 56; Gaps 17;

QY 4 ASKAPQSGAGSAGQLVPE-----VNTADPLPMPVAGPTTAVATAGQVNMIDPWVNN 59
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 151 AESDGP-----GSAEIVTEEQGVVQQQPAP-APTALATLATATGK-SVEQEWMT-- 199
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 60 FVQSPQSEFTTSPNNTPGDILFDLQGLHNPFLSHLSQMYNGWGNMVRILLAGNAPS 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 200 -FFSYHTSINWSTVBSQKILYSQALNPSINPYLDHIAKLYSTWGGIDVRFVTSQGVF 258
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 AGKIIVCCVPBGFTS-SSLTIAQATLFPHVIAADVTLERPLEDMVRNVLVHTNDNQPT 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 259 GGLKALILVPPGVPIESVSMLQ-----YPHLVFDARQTEPVITPIRKTLYHSMDDTTTLRLVIWYNEL 315
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 179 MRLVCLYTLPTLTGSGNSDFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIP 238
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 316 TKLV---INPYENGVENKTSITVE---TFPSADTFALLKPPGSLIKGSIPLSLIP 368
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 239 LQTL---SNSRPSLLIQMILSPDASQVVOFQNGRCL-IDGQLLG-TTPATSGQLFRVRGK 294
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 369 RNSAHMNGRNMWSTISGFSVQPRV-----FQSNRHFDFDSTTTGWSVTPYVPIEIKIQK 423
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 295 INQARTLNLTEDGKPFMAEDSPAPVGFDPFGKCDHMRISKTNNNTSSGDMRMSVQ 354
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 424 VGSNNKWFHVIDTD-----KALVPGIGDGPDDTTIPD---ETKATNGNFSYGESYAGST- 475
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 355 TNVQGVFHLGSIQFDEVFNHPTGDYI-GTIEWISQ 390
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 476 -----TIKENENSTHFKGYICGTLSTEIP 501
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 10
A48562
coat protein - San Miguel sea lion virus (serotype 1)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: A48562
R:Neill, J.D.
Virus Res. 24, 211-222, 1992
A:Title: Nucleotide sequence of the capsid protein gene of San Miguel sea
eins.
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Accession: A48562
A:Molecule type: genomic RNA
A:Residues: 1-702 <NEI>
A:Cross-references: GB:M87481; NID:g334882; PIDN:AAAL6217.1; PID:g334884
A:Note: sequence extracted from NCBI backbone (NCBIN:113564, NCBIP:113565)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.1%; Score 293.5; DB 1; Length 702;
Best Local Similarity 28.3%; Pred. No. 7.4e-14;
Matches 102; Conservative 45; Mismatches 178; Indels 35; Gaps 12;

QY 12 ADGASGAGQLVPEVNTADPLPMPVAGPTTAVATAGQVNMID-PWIVNVFVQSPQSEFTI 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 153 SDGPGGADIVTEEQGVVQQQPAPQASALTTLAAASTGKTVDCEWTT---FFSYHTAVNM 209
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 71 SPNTPGDIILFDLQGLHNPFLSHLSQMYNGWGNMVRILLAGNAPSAGKIIVCCVPP 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 210 STTEAQQKILFSRALSPELNPLYRHSLSYSGGIDVRFVTSQGVFGVGGKLAALIVPP 269
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 131 GFTS-SSLTIAQATLFPHVIAADVTLERPLEDMVRNVLVHTNDNQPTMLVCLMXTPL 189
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 270 GIEPVESFTMLQ---YPHLVFDARQTEPVITPIRKTLYHSMDDTTTLRLVIWYNEL 326
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 190 RTGGSGNSDFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIPLOT--LSNSRF 247
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 327 INPVQSBPKS-SCSITVETRPSSDFTFSLKPPGSLIKGSIPLSRNRHWMGNRW 385
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 248 PSLIQMILSPDASQVVOFQNGRCL-IDGQLLGTTPATSGQLFRVRGKINQARTLNUTE 306
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 386 WSTDGFFVQPRV-----FQSNRHFDFDSTTTGWSVTPYIPIEVLTLEKLDRGQVFKVTD 440
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 307 VDGKPFMAEDSPAPVGFDPFGKCDHMRISKTNNNTSSGD-----PMRSVSVQTNVQGF 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 441 TE-----KSLVP-GLPD-----GWPDTTIPMTASNGNDYTVAEYRITNNGTHFKGF 488
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
VCWVFC
coat protein - feline calicivirus (strain Japanese F4)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C:Accession: B40481
R:Toyba, Y.; Taniguchi, Y.; Takahashi, E.; Utagawa, E.; Takeda, N.; Yamazaki, K.; Yamazaki, K.
Virology 183, 810-814, 1991
A:Title: Sequence analysis of the 3'-end of feline calicivirus genome.
A:Reference number: A40481; MUID:91306470; PMID:1853578
A:Accession: B40481
A:Molecule type: genomic RNA
A:Residues: 1-668 <TOH>
A:Cross-references: GB:D90357; NID:g221264; PIDN:BAAL4371.1; PID:g221266
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,301,304,399,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.8%; Score 284; DB 1; Length 668;
Best Local Similarity 24.7%; Pred. No. 3.5e-13;
Matches 145; Conservative 74; Mismatches 246; Indels 122; Gaps 27;
```

Qy	23	PEVNTA-DPLPMEPVAGETTAVATAGVNMIDPIMVNNVQSPQGEFTI-----SPNPT	75
Db	133	PEOCTAVGGVIAEESAQMSTMAASGKGSVSEW-----EAFSFHTSVNMWSTSET	183
Qy	76	PGDILFDLQGLPHNPPFLSHLSOMYGVGNMVRVILLAGNAGSAGKIIVCVCPGPGFTS-	134
Db	184	QGKILFKQSLGPLNPLYLEHLSKLVAMSGSIEVRFSGSGVFCGKLAALVVPFGVDVP	243
Qy	135	SSLTAQATLPPHYIADVRLTEPIEMPLEDVYRNLYHTNDNQPMRLVCMLYTPL-RTGG	193
Db	244	QSTSMLO--YPHVLFDARQVEPVITPDLRSLTYLHVMSDDTTLSLIMVYNDLINEPYA	300
Qy	194	SGNSDSFVAVAGRVLTAPSSDFSLFLVPPPTIEQKTRAFVNPINPLQTLSNS-----	246
Db	301	NDSNSSGGCIIV--VETKEPGDFKFHLLKPP-----GSVLTHGSIIPSDILPKSSSLWIGNR	353
Qy	247	FPSLIQGMILSPDASQVVQFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQOARTLNITE	306
Db	354	IYWDITDFVIRP-----VFVQANRHFDNQ--ETAGWSTPRFR-----PITIITISE	397
Qy	307	VDGKPF---MAFD---SPAPVGGDFGKCDWHMRIKSTPNNTSGDPMRS-----V	351
Db	398	KNGSKLGIQVATDYIIPGIPDGWPDTTIADKLI PAGDYSITTGEGNDIKTAQAYDTAAVV	457
Qy	352	SVQTNVQGFVHLGSIQDFEVNHPETGD-YIGTIEWISQP-----STPGTDINLWEIPD	405
Db	458	KNTTNFRGWY-ICGSLQ-----RAWGDKKISNTAFITTAIRDGNEIKPSNITIDMTKLAIV	510
Qy	406	Y-GSSLSQAANLAPVFPF---PGFGEAL-----VYFVSAPFPNNRSA-----PND	447
Db	511	YQDTHVEQVQTSDDTALLGYTGIGEEAIGNRDRVRVVISVLPAGAGGNHPFYKNS	570
Qy	448	VPCLLPQEIYTHFVSEQAPTMGDAALLHVDP-----DTRNRLGEFKLYPGGYLT	497
Db	571	IKLGIVIRSIDVFNSQLHTSRQLSINHYLLPDPSPFAVYRIIDNSGWFIDIGIDSEGF--	628
Qy	498	CVPNGVGA---GPQOLPLNGVPLFVSWVSRSFVQLKPVGVGTASTARSL	541
Db	629	---SFVGSYSDIGKLBFFPLSASVYGI-----QLAKIRLASHIRSGM	665

RESULT 12
JQ2354
capsid protein - feline calicivirus (strain NADC)
C/Species: feline calicivirus
C/Date: 28-Aug-1995 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C/Accession: JQ2354
F:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A/Title: Analysis of feline calicivirus capsid protein genes: identification of variable
A/Reference number: JQ2354; MUID:94065683; PMID:7504075

Query Match	9.5%	Score 275.5;	DB 2;	Length 668;
Best Local Similarity	25.1%	Pred. No. 1.5e-12;		
Matches 145;	Conservative 70;	Mismatches 236;	Indels 127;	Gaps 28;
QY	34	EPVAGPTTAVATAGQVNMIDPMVNVFVQS PQGEFTI-----SENNTPGDILFDLQGLP	87	
DB	145	EPSAQMTAAADMATCKSDSEW-----EAFSFHTSVNMSTSETQGKILFKQSLGP	195	
QY	88	HLNPFLSHLSOMYNGWGNMVRILLAGNAFSAKGKIIIVCCVPBPGETS--SSLTIAQATLFP	146	
DB	196	LLNPFLSHLSKLYVWMSGSEVRFISGSGVGGKLAALIVPPGVDVPQISIMLQ---YP	252	
QY	147	HVIADVRLTEPIEMPLEDVRNLYLHTNDNQPTMRLVCMLYTPL-RTGGSGNSDSFFVAG	205	
DB	253	HVLFARQVDVPFISGPLRSLTSLHLMPTDITTSIVMWNDLINPYANDSNSSGCIYV--	311	

Qy	206	RVLTA	SSDPSFLFVPP	PTIEQKTRAF	TVPNILP	LTLSN	-----R	FP	SLIQ	GM	ILSP	250
Db	312	-VET	KGPDPK	FHLLKPP	-----G	SML	THG	S	V	P	D	365
Qy	259	DA	QVQFQNGR	CLIDQ	LLGTTP	ATSP	SGQ	FR	VRG	KINQ	ARTLN	315
Db	366	-----F	VQ	ANR	FN	FNQ	-----E	T	AG	S	AP	409
Qy	316	DS	PAPVGF	PD	FGK	D	WHM	R	SK	TP	-----N	361
Db	410	DY	IVP	-G	IPD	-G	W	P	D	T	T	467
Qy	362	PH	GS	IQ	F	E	V	N	H	P	T	410
Db	468	-I	CS	LQ	-----R	AWG	K	K	I	S	T	520
Qy	411	SQ	AN	L	A	P	P	P	P	P	F	457
Db	521	Q	T	S	D	A	L	A	L	L	G	580
Qy	458	TH	F	S	E	O	A	P	T	M	G	504
Db	581	D	V	F	N	S	Q	L	T	H	S	634
Qy	505	-A	G	P	O	O	L	P	L	N	G	541
Db	635	S	I	G	K	E	V	P	L	S	A	665

RESULT 13
 JQ2356
 capsid protein - feline calicivirus (strain KCD)
 C:Species: feline calicivirus
 C:Date: 28-Aug-1995 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
 C:Accession: JQ2356
 R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.
 J. Gen. Virol. 74, 2519-2524, 1993
 A:Title: Analysis of feline calicivirus capsid protein genes: identification of
 A:Reference number: JQ2354; MUID:94065683; PMID:7504075
 A:Accession: JQ2356
 A:Molecule type: mRNA
 A:Residues: 1-668 <SRA>
 A:Cross-references: GB:L09719; NID:G305107; PID:AA16487.1; PID:G305108
 C:Superfamily: feline calicivirus coat protein
 C:Keywords: capsid protein

Query Match	9.5%	Score	275.5	DB 2	Length	668			
Best Local Similarity	26.0%	Pred. No.	1.5e-12						
Matches	125	Conservative	58	Mismatches	167	Indels	131	Gaps	24
QY	34	EPVAGPTTAVATAGQVNMIDPWINVNFVQSPQGEFTI	-----SNNTFGDILFDQLQLGP	87					
			:::	:::	:::	:::	:::	:::	:::
Db	145	EPNAQMSTAADWATCKSDVSEW	-----EAFSPHTSVNMWSTSETQKILFKQSLCP	195					
			:::	:::	:::	:::	:::	:::	:::
QY	88	HLNPFLSHLSQMYNGWGNVRRILLAGNAFSAGKIIVCCVPPGGTIS	---SSLTIAQATLFP	146					
			:::	:::	:::	:::	:::	:::	:::
Db	196	LINPYLEHLAKLYAWMSGIEVRFISGSGVFGGKLAIVVPPGVDVQOSTSMLQ	---YP	252					
			:::	:::	:::	:::	:::	:::	:::
QY	147	HVIADVRLTEPIEMPLEDVRNVLYHTNDNQPTMLRVLCMLYPL	-RTGGSGNSDSFFVAG	205					
			:::	:::	:::	:::	:::	:::	:::
Db	253	HVLFDAQVEPIFISPDRLSTLYHLMSTDYTTSLVWAYNDLINPYANDSNSGCIVT	-	311					
			:::	:::	:::	:::	:::	:::	:::
QY	206	RVLTAPSDDFGFLFLVPP	-----TIEQKTRAFYTPVNPILQTLNSRFPSLIQM	254					
			:::	:::	:::	:::	:::	:::	:::
Db	312	-VETKPGSDFRPHLLKPPGSLVTHGVSFSDLPKPTSSLMIGN	-----RYWNTITNF	361					
			:::	:::	:::	:::	:::	:::	:::
QY	255	ILSPDASQVQFQNGRCILIDGOLLGTTATSQGLFRVGRGKINQAGARTLNLTEVDGKPF	--	312					
		:::	:::	:::	:::	:::	:::	:::	:::
Db	362	VIKP-----FVFOAHRHFDENQ	---ETAGWSTPRFR	405					
			:::	:::	:::	:::	:::	:::	:::
QY	313	-MAFDSPAPVGFPPDFGKCDWHMRISKPTNPNTSSGDPMRSSVQTVNQGVFPHLGSITQFDE	371						
			:::	:::	:::	:::	:::	:::	:::

Db 406 GVATDYIVP-GIPD-----GW-----PDTTIS-----EE 428
QY 372 VFNHPTGDIYGTIEWISQSTPTPGTD-----INLWEIPDYSSLSQAANLAPVFPFP 423
Db 429 LI--PAGDYAITNDIGNDITTPAGYDAADTKNNTNFRGMVCG-SLQRA-----475
QY 424 GESEALV---YFVSAPFGNNRSAPNDVCLLPQEVITHF---VSEQAPTWGDA-ALLH 475
Db 476 -WGDKKISNTAFITATVKNKLPKSN---MIDQTKIAVFDQNHGVGDKVQTSDDTLALIG 531
QY 476 Y 476
Db 532 Y 532

RESULT 14

VCWVFF
coat protein - feline calicivirus (strain F9)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C:Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text_change 28-Jul-2000
C:Accession: B43382; C45538; PQ0407; S23702
R:Carver, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.;
Virology 190, 443-448, 1992
A:Title: The complete nucleotide sequence of a feline calicivirus.
A:Reference number: A43382; MUID:92410623; PMID:1529544
A:Accession: B43382
A:Molecule type: Genomic RNA
A:Residues: 1-671 <CARI>
A:Cross-references: GB:M86379; NID:G323877; PIDN:AAA79327.1; PID:G323879
R:Carver, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.
Arch. Virol. 122, 223-235, 1992
A:Title: Identification and sequence determination of the capsid protein gene of feline
A:Reference number: A45538; MUID:92117861; PMID:1731695
A:Accession: C45538
A:Molecule type: genomic RNA; protein
A:Residues: 1-671 <CAR>
A:Cross-references: GB:M86379; NID:G323877; PIDN:AAA79327.1; PID:G323879
A:Experimental source: strain F9
A:Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)
R:Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.
J. Gen. Virol. 73, 2429-2433, 1992
A:Title: The cloning, sequencing and expression of a major antigenic region from the fel
A:Reference number: PQ0407; MUID:93019069; PMID:1402818
A:Accession: PQ0407
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 217-266, 'SI', 269-336, 'V', 338-395, 'S', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>
A:Cross-references: PIDN:AB23553.1; PID:G257083
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,304,439,618/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.3%; Score 268.5; DB 1; Length 671;
Best Local Similarity 26.8%; Pred. No. 5.1e-12;
Matches 110; Conservative 47; Mismatches 149; Indels 105; Gaps 18;
QY 10 QSADGASGAGQLVPEVNT-ADPLPEPVAGPTTAVATAGQVNMIDPWVNNFVQSPQGBF 68
Db 124 EADGSITH-----PQGTWVGVIAPSAQMTAADMATGKSVDSW-----EAPF 170
QY 69 TI-----SPNNTPGDILFDLQGLHNPFLSHLSQMYNGWGNMVRILLAGNAFSAK 122
Db 171 SPHTSVNWSSTETQKILFKQSLGPLLPYLEHLAKLYVAMSSIEVRFSGSGVFGCK 230
QY 123 IIVCCVPPGFTS-SSLTIAQATLPHVIAVTLSPLEPIEDVRNLYHTNDNQPTMRL 181
Db 231 LAIVVPPGVDPVQSTMLQ---YHVLFDARQVEVPFICLPDLRSTLYHLSMDTDTTSL 287
QY 182 VCMLYTFL-RTGGSGNSDSFVAVAGVLTAPSSDSFSLFVPP-----TIEQKT 229
Db 288 VIMVYNDLINPYANDANSGGCIVT--VETKPGDPDFKHLKPPGSMLTGHSIPSLIPKT 345

QY 230 RAFTVNPITLQTLNSRFPSSLIOGMILSPDASQVVFQNGRCGLIDCOLLGTTPATSGQLF 289
Db 346 SSLWIGN-----RYWSDITDFVIRP-----FVQANRHDFNQ-----ETAGWSTPRF 308
QY 290 RYRGKINGQARTLNLTEVDGKFP---MAFDSAPVGVFPDFGKCDMHMRISKTPNNTSSGD 346
Db 389 R-----PISVTITEQNGAKLGGVATDYIVP-GIPD-----GW-----PDTTIPGE 428
QY 347 PMRSVSVQTNVQGVFPHLGSIQFDEVNHPHTQDYIGTIEWISQSTPTPGTD 397
Db 429 LI-----PAGDYAITNGTGNDITATGYD 452

RESULT 15

VCWVFF
coat protein - feline calicivirus (strain CFI/68 FIV)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C:Date: 30-Jun-1992 #sequence revision 30-Jun-1992 #text_change 20-Aug-1999
C:Accession: A40507; B40507; T09246
R:Neill, J.D.; Reardon, I.M.; Heinrichson, R.L.
J. Virol. 65, 5440-5447, 1991
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline calici
A:Reference number: A40507; MUID:91374597; PMID:1716692
A:Accession: A40507
A:Molecule type: genomic RNA
A:Residues: 1-668 <NEI>
A:Cross-references: GB:M32819; NID:G323874; PIDN:AAA42925.1; PID:G323875
A:Accession: B40507
A:Molecule type: protein
A:Residues: 373-379; 403-419; 481-489; 560-566 <NE2>
R:Neill, J.D.
submitted to the EMBL Data Library, April 1998
A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.
A:Reference number: Z16626
A:Accession: T09246
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: genomic RNA
A:Residues: 1-668 <NE3>
A:Cross-references: EMBL:U13992; NID:G3056875; PIDN:AAC13993.1; PID:G537256
A:Experimental source: strain CFI/68 FIV
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.2%; Score 267; DB 1; Length 668;
Best Local Similarity 31.7%; Pred. No. 6.5e-12;
Matches 76; Conservative 33; Mismatches 97; Indels 34; Gaps 8;
QY 34 EPVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEFTI-----SPNNTPGDILFDLQGLP 87
Db 145 EPNAQMTAADMATGKSVDSW-----EAPFSPTSVNWSSTETQKILFKQSLP 195
QY 88 HLNPLSHLSQMYNGWGNMVRILLAGNAFSAKIIIVCCVPPGFTS-SSLTIAQATLPP 146
Db 196 LLNPYLTHLAKLYVAMSGSDVRFSGSGVFGKLAALIVVPPGIDPVQSTMLQ---Y 252
QY 147 HVIAVDRTLEPTMLEPDRVNLVHTNDNQPTMRLVCMLYTFL-RTGGSGNSDSFVAVAG 205
Db 253 HVLFDPARQVEVPFIPDLRSTLYHLSMDTDTTSLVIMVYNDLINPYANDANSGGCIVT- 311
QY 206 RVLTAAPSSDSFSLFVAPPTIEQKTRAFTVNPITLQTLNS-----RPPSLIOGMILSP 258
Db 312 -VETKPGDPDFKHLKPP-----GSMLTGHSIPSLIPKSSSLWIGNRFSWSDITDFVIRP 365

Search completed: June 1, 2004, 13:55:20
Job time : 12.5352 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:36:02 ; Search time 10.4195 Seconds
(without alignments)
4985.230 Million cell updates/sec

Title: US-09-926-799-7
Perfect score: 2845
Sequence: 1 MKMASNDATPSNDGAAGLV.....VNQFYSLAPMTGNGRRRVQ 540

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2339	82.2	542	2 S60616	capsid protein - h
2	1870.5	65.7	539	2 S40111	capsid protein - h
3	1229	43.2	530	2 B37471	capsid protein - N
4	1168.5	41.1	546	2 B37491	major capsid prote
5	328	11.5	2344	2 S53399	genome polyprotein
6	320.5	11.3	2344	1 RRWRH	genome polyprotein
7	318	11.2	702	1 A48562	coat protein - San
8	316	11.1	2344	2 S64740	genome polyprotein
9	313	11.0	576	2 A53982	capsid protein - E
10	292.5	10.3	668	1 VCMWF	coat protein - fel
11	292.5	10.3	671	1 VCMWF9	coat protein - fel
12	283.5	10.0	668	2 JQ2354	capsid protein - f
13	280.5	9.9	668	1 VCMWF	coat protein - fel
14	274	9.6	703	1 C48562	coat protein - San
15	269	9.5	668	2 JQ2356	capsid protein - f
16	138	4.9	2333	1 GNNY2F	genome polyprotein
17	137	4.8	2332	1 GNNYF	genome polyprotein
18	136	4.8	1011	1 GNNYC1	genome polyprotein
19	131.5	4.6	733	2 JQ1891	capsid protein - f
20	131	4.6	733	2 JQ1892	capsid protein - f
21	129	4.5	2336	2 S37077	genome polyprotein
22	128.5	4.5	2207	2 S09553	genome polyprotein
23	128	4.5	3624	2 AD0835	large repetitive p
24	127.5	4.5	416	2 T50279	hypothetical serin
25	127	4.5	2332	1 GNNY4F	genome polyprotein
26	126.5	4.4	1978	2 S77257	hypothetical prote
27	124.5	4.4	2205	1 GNNY2W	genome polyprotein
28	123	4.3	2194	1 GNNY7	genome polyprotein
29	121.5	4.3	3085	2 T00327	polyprotein - infe

30 120.5 4.2 2207 1 GNNY5P genome polyprotein
31 119 4.2 757 2 JS0198 genome polyprotein
32 118.5 4.2 2206 1 GNNY27 genome polyprotein
33 118 4.1 2209 1 GNNY2P genome polyprotein
34 117.5 4.1 5376 2 T42215 zonadhesin - mouse
35 115.5 4.1 2175 1 GNNYBE genome polyprotein
36 114.5 4.0 2206 2 S03822 genome polyprotein
37 114.5 4.0 6658 2 T13931 projectin - fruit
38 114 4.0 2207 1 GNNY1P genome polyprotein
39 114 4.0 4199 2 S76412 hypothetical prote
40 113 4.0 2185 1 GNNYSV genome polyprotein
41 112.5 4.0 925 2 S50490 hypothetical prote
42 112.5 4.0 2206 1 GNNY4P genome polyprotein
43 112 3.9 2209 1 GNNY3P genome polyprotein
44 109 3.8 257 2 AH0859 periplasmic fibril
45 108.5 3.8 940 2 A40985 projectin - fruit

ALIGNMENTS

RESULT 1
S60616
capsid protein - human calicivirus (strain Melksham)
C:Species: human calicivirus
A:Variety: strain Melksham
C:Date: 23-May-1997 #sequence_revision 23-May-1997 #text_change 28-Jul-2000
C:Accession: S60616
R:Green, S.M.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
Virus Res. 37, 271-283, 1995
A:Title: capsid protein of a group 1 round-structured virus: molecular characterization
A:Reference numbers: S60616, U00006036658, F0440859362
A:Accession: S60616
A:Molecule type: genomic RNA
A:Residues: 1-542 <GRE>
A:Cross-references: EMBL:X81879; NID:G976077; PID:CAA57462.1; PID:G976079
A:Experimental source: strain Melksham
A>Note: it is uncertain whether Met-1 or Met-3 is the initiator
C:Superfamily: human calicivirus capsid protein
C:Keywords: capsid protein; coat protein

Query Match 82.2%; Score 2339; DB 2; Length 542;
Best Local Similarity 78.9%; Pred. No. 4.2e-164;
Matches 430; Conservative 56; Mismatches 51; Indels 8; Gaps 3;

QY 1 MKMASNDATPSNDGAAGLVPSNNEAMALEPVVGASLAAPVTGQTNIDPWIRTNFVOAP 60
Db 1 MKMASNDAPSTDGAGLVPSNNEAMALEPVVGASLAAPVTGQTNIDPWIRTNFVOAP 60

QY 61 NGEFTVSPRNGSGEILVNLELGPENPYLAHLARMYNGYAGGMEVQVMLAGNAFTAGKI 120
Db 61 NGEFTVSPRNGSGEILVNLELGPENPYLAHLARMYNGYAGGMEVQVMLAGNAFTAGKI 120

QY 121 FAAPPPYFVENLSQITMFPVLIIDVRLPVLPPDVRSITLPHNQKDEPKMLVA 180
Db 121 FAAPPPYFVENLSQITMFPVLIIDVRLPVLPPDVRSITLPHNQKDEPKMLVA 180

QY 181 MLYTLRNSGDDVFTVSCRLLTRPSPEFDTYLVPPVTESKTPFTLPVLTGLSELS 240
Db 181 MLYTLRNSGDDVFTVSCRLLTRPSPEFDTYLVPPVTESKTPFTLPVLTGLSELS 240

QY 241 RFLPSIDEMVTPNSISVVQPNQGRVTLDELGLGTQLQACNICSIRGKVTQVPSQHM 300
Db 241 RFPVFDQMYTSPNEVISVQCNQGRCTLDGELQGTQLQVSGICAPKGEVTAHLHDNDL 300

QY 301 WNLEITNLNQTFDDTDVAPLGPDPFAGEVGVLSQNR----GSPNAPRAHDVVA 356
Db 301 NNVTITNLNSPFPDSEDIPAPLGPDPFAGEVGVLSQNR----GSPNAPRAHDVVA 356

QY 357 TVSDKYTPKLGIVQIGTWTNDVE--NQPTKFTPIGLNEVANGHRFEQMTLPYSGALTIN 415
Db 361 TYTAQTPKLGIVQIGTWTNDVE--NQPTKFTPIGLNEVANGHRFEQMTLPYSGALTIN 415

Db 537 R 537

RESULT 3

B37471

capsid protein - Norwalk virus

C:Species: Norwalk virus

C:Date: 24-Feb-1994 #sequence_revision 25-Apr-1997 #text_change 28-Jul-2000

C:Accession: B37471

R:Jiang, X.; Wang, M.; Wang, K.; Bates, M. K.

R:Jiang, X.; Wang, M.; Wang, K.; Bates, M. K.

R:Jiang, X.; Wang, M.; Wang, K.; Bates, M. K.

A:Title: Sequence and genomic organization of Norwalk virus.

A:Cross-References: GB:M87661; NID:g1061311; PIDN:AAB50466.1; PID:g1061313

A:Note: sequence extracted from NCBI backbone (NCBIP:134157)

A:Superfamily: human calicivirus capsid protein

A:Accession: B37471

A:Status: Preliminary; not compared with conceptual translation

A:Molecule type: genomic RNA

A:Residues: 1-530 <GJA>

A:Cross-References: GB:M87661; NID:g1061311; PIDN:AAB50466.1; PID:g1061313

A:Note: sequence extracted from NCBI backbone (NCBIP:134157)

A:Superfamily: human calicivirus capsid protein

Query Match 43.2%; Score 1229; DB 2; Length 530;

Best Local Similarity 47.4%; Pred. No. 1.6e-82;

Matches 262; Conservative 82; Mismatches 163; Indels 46; Gaps 11;

QY 1 MKMASNDATPSNDGAG--LVPSN-NEAMALEPFWGASLAAPVTGQTNIIDPIRTNF 56

DB 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMPDVAGSTAVATAGQVNPIDPWIINN 60

QY 57 VOAPNGEFTVSPRSPGELVNLGPELNPYLALHARMYNGAGGMEVQVMLAGNAFTA 116

DB 61 VOAPNGEFTVSPRSPGELVNLGPELNPYLALHARMYNGAGGMEVQVMLAGNAFTA 120

QY 117 GKIIIAAAPPYPPVENLSPSQITMPPHVIIDVRLPVLLEPMPDVRSTLPHFNQKDEPKM 176

DB 121 GKIIIVSCIPPGSGSHLTIQAATLFPVIAADVRLDPIEVLDPVNRVLFHNNDRNQOTM 180

QY 177 RLVMALYTPLRNSGGDDVFTVSCRILTRPSPEFTYLVPTVESKTKPTLPLVTIGE 236

DB 181 RLVMALYTPLRNSGGDDVFTVSCRILTRPSPEFTYLVPTVESKTKPTLPLVTIGE 240

QY 237 LNSRPFSLIDEMWTPSNESIVVQPNQGRVTLDDGELLGTQLOACNICIRGKVTGVPS 296

DB 241 LNSRAPLPSISIGISPDNVQSVQNGRCTLGRLVGTTPVSLSHVAKIRGTSNGTV-- 298

QY 297 ECHMNLEITNLNGTQDFTDDVPAPLGVDPFAG-----EVFGVLQSRNGESNPANR 349

DB 299 -----INLTLDGTPPHFEG-PAPIGPDLGGCDWHINMTQFHSQTQ----- 342

QY 350 AHDVAVTYSKYTKPLGLVQ--TGTWNTNDVENQPTKFTPIGLNEVANGHRFQWTL 406

DB 343 -YD--YDTPDFTFVPHLGSIQANGISGNYGVG-----LSWISPSHPSGQVDLWKIP 393

QY 407 RYSGALTNNMLAPAVAPLPGERLLFFRSYVPLKGGFGNPAIDCSVPQEWVQHIFYQESA 466

DB 394 NYGSSITEATHLAPSVYPPGFEVLVFFWMSKMPGAYNLP---CLLPQEYIASHLASEQA 450

QY 467 PSLGDVALRYNPDTGRVLFKAHLKGGGLTV--SSSTGTVVVPVANGYFKFDSVNVQF 524

DB 451 PTVGEAALLHYVDPTGRNLGEFFKAYPDGFLTCVPNGASSGQQQLPINFVGVFVSVSRF 510

QY 525 YSLAPMGTCNGRR 537

DB 511 YQLKPVGTASSAR 523

RESULT 4

B37491

major capsid protein [similarity] - Southampton virus

N:Alternate names: orf2 protein

C:Species: Southampton virus

C:Date: 22-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 18-Aug-2000

C:Accession: B37491
R:Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
Science 259, 516-519, 1993
A:Title: Sequence and genome organization of a human small round-structured (Norwalk-like
A:Reference number: A37491; MUID:93142023; PMID:8380940
A:Accession: B37491
A>Status: not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-546 <LAM>
A:Cross-references: GB:L07418; NID:g1236787; PIDN:AAA92984.1; PID:g295114
A>Note: sequence extracted from NCBI backbone (NCBIP:123458)
A:Note: small round-structured virus, SRSV, Norwalk virus, Norwalk-like virus, serotype:
C:Superfamily: human calicivirus capsid protein
C:Keywords: glycoprotein
F:303,340,441/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.1%; Score 1168.5; DB 2; Length 546;
Best Local Similarity 44.9%; Pred. No. 4.8e-78;
Matches 254; Conservative 86; Mismatches 179; Indels 47; Gaps 13;
QY 1 MMASNDATPSNDGAG---LVPESNN-EAMALEPVGASLAAPVTGQTNIIDPWIRTF 56
DB 1 MMASKADPQSDAGSAGQALPEVNTADPLPMEFVAGPTTAVATAGQVNMIDPWVNF 60
QY 57 VQAPNGEFTVSPRNGEILLNLELGPENLPIYLAHMYNGYAGMEVQVMAGNAFTA 116
DB 61 VQSQOGEFTISPNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGNMVRILLAGNAFSA 120
QY 117 GKIIIAAAPPYPVENLSPQITMPHVIIDVRLPEVLLPMDVRSITLPHENQKDEPKM 176
DB 121 GKIIIVCCVPPGTSSTLTAAQLPFPVADVRLTEIEMPLEDRVNLVHYTND-NQPTM 179
QY 177 RLVAAMLYPLRS-NGSGD-DVFTVSCRILTRSPSEFDFLYLPPTVESKTEFTLPLVLT 234
DB 180 RLVCMLYPLRTGGSGNSDSFVAGRVLTAFSSDFSLFVLPPIEQKTRAFVTPNPL 239
QY 235 GELNSRPLSIDEMVTSNBSIIVQPNQGRVTLGELLGTTLOACNICSIRGKVTGV 294
DB 240 QTLNSRPLSQMILSPDASQVQFQNGRCLIDQLLGTTPATSGQLFRVRGKI---- 295
QY 295 PSEOHMNLIEITNLNGTDPDPPAPLGVDPFAGEVFGVLSQRNRESNANRAHDAV 354
DB 296 --NQARTLNLTEVDKPKMAFDS-PAPVGFDFGKCDMHMIRISKTPNNTGSGDPMRSVS 352
QY 355 VATYSDKYTPKLGVLQIGTWNVDENOPT-----KFTPIGLNEVANGHRPEQ 402
DB 353 VQTNVQGFPHLGSIQF-----DEVNHTGDIYGTIEWISQSPTPPGTD-----INL 400
QY 403 WTLPRYSGALTLMNLAPAVLFPGERLLFRSVPLKGGFNP-AIDCSVPQBWQHF 461
DB 401 WEIPDYGSLQAANLAPPVPPGFEALVYFVSAPFGPNRSAPNDVPCLLPQEVYTHF 460
QY 462 YQESARSLGDVALVYVNPDTGRVLEAKLHGGELTV--SSTSTGTPVVVANGYFKDS 519
DB 461 VSEQAPTGMGDAALLHYVDPTNRNLGFEKLYPGGYLTCVPGNGVAGPQOPLPLNGVLFVS 520
QY 520 WYNQFVSLAPMGTGN-----GRRRV 539
DB 521 WVSRYQLKPVGTASTARGLVRII 546

RESULT 5

S55399
genome polypeptide - rabbit hemorrhagic disease virus (isolate BS89)
A:Species: rabbit hemorrhagic disease virus
A:Variety: isolate BS89
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 18-Jun-1999
C:Accession: S55399
R:Rossi, C.
submitted to the EMBL Data Library, May 1995
A:Reference number: S55399
A:Accession: S55399
A:Molecule type: genomic RNA

A:Residues: 1-2344 <ROS>
A:Cross-references: EMBL:X87607; NID:g854640; PIDN:CAA60910.1; PID:g854641
A:Experimental source: isolate BS89
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: polypeptide

Query Match 11.5%; Score 328; DB 2; Length 2344;
Best Local Similarity 26.1%; Pred. No. 2.6e-15;
Matches 155; Conservative 77; Mismatches 210; Indels 152; Gaps 30;
QY 3 MASNDATPSNDGAGL-----VPSNNEAMALEPVV-----CASLAAPVTG-- 43
DB 1766 MEKARTAPOGAGTATATASVPGTTD--GLDPGVATTSVVTAENSSASATATAGIGP 1823
QY 44 --QTNIIDPWIRTFVQAPNGEFTVSPRNSPGEILLNLELGPENLPIYLAHMYNGYAG 101
DB 1824 PQOVDOQETW-RTNPFY--NDVFTSVADAPGSIITYVQHSQNNPFTAVLSQMVAGWAG 1880
QY 102 GHEVQVMLAGNAFTAGKIIFAAVPPYFPVENLSPS-QITMPHVIIDVRLPEVLLPMD 160
DB 1891 GMQFRFIVAGSVFGGRLLVAIVPPGI---EIGPGLVRFQFPHVVIDARSLEVTITMPD 1937
QY 161 VRSTLPHFNQKDEPKM--RLVAMLYPLRSNGSGDDVFTVSCRILTRSPSEDFTVLVP 218
DB 1938 LRPNMYH--PTGDPGLVPTLVLSVYNNL-INPFGSTSAIQVTVTRPSEDFEFVIRAP 1994
QY 219 ---TVESKTKP--FTLPVLTGELNSRPLSIDEMVTSNBSIV-VQPNQGRVT----- 267
DB 1995 SSKTVDSISPAGLLTTPVLT-GVGNDNRW-----NGQIVGLQVPVGGFSTCNRH 2042
QY 268 --LDGELLGTTLOACNICSIRGKVTGVQVPSQHMNLNLEITNLNGTDPDTPDVPAPLV 325
DB 2043 WNLGSTGYGWSPPFADIDHRRGSASYPGNSATNVLQFWYANAGSAVDNPISQV-APDGF 2101
QY 326 PDF-----AGEV-FGLVLSQRNRESNAN-RAHDAAVAT----- 357
DB 2102 PMSFVPPNGPCI PAAGWVGFAIWNSSGAPNVTTVQVLELGFATGAPGNLOPTNTSG 2161
QY 358 ----YSDKYTPKLGVLQ-----IGTWNVDENOPT----- 384
DB 2162 AQTVAKSIYAVVTGTAQNPAGLFVWASGVISTPNANAITTTPQDRIVTTTCTPAAAPVG 2221
QY 385 KFTPIGLNEV-----ANGHRPEQWTLPRYSGALTLMNLAPAVLFPGERL 431
DB 2222 KNTPIMFASVVRRTGDVNNATAGSANGTQGTGSQLP---PVTIGLSNLYSSALMPGQ-- 2276
QY 432 LFERSYVPLKGGFNGPAIDCSVPQEWQHFQESAPS-----LGDVALVRYVNP 480
DB 2277 --FFVWQLTFASGF--MEIGLSVD-----GYFAGTAGSTTLIDLTELIDVRPVG 2323

RESULT 6

RKWRH
genome polypeptide - rabbit hemorrhagic disease virus
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C:Species: rabbit hemorrhagic disease virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 29-May-1998
C:Accession: A41039
R:Meyers, G.; Wirblich, C.; Thiel, H.J.
Virology 184, 664-676, 1991
A:Title: Rabbit hemorrhagic disease virus--molecular cloning and nucleotide sequencing
A:Reference number: A41039; MUID:91361557; PMID:1840711
A:Accession: A41039
A:Molecule type: genomic RNA
A:Residues: 1-2344 <MEY>
A:Cross-references: GB:M67473
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: nucleotidyltransferase

Query Match 11.3%; Score 320.5; DB 1; Length 2344;
Best Local Similarity 25.9%; Pred. No. 9.3e-15;
Matches 152; Conservative 78; Mismatches 209; Indels 147; Gaps 30;

QY 8 ATPSDNGA---AGLVPSNNEM-----ALEPVV-----GASLAAPVTG-----QTNIID 49
 Db 1772 AAPGEAAGATTASVPGTTDGMDPGVVATTSVITAENSSASITAGIGPPQOVQOE 1831
 QY 50 PWIRNFVQAPNGETVSPRNSPGILLVNLGLPELNPVLAHLARMYNGYAGGMEVQVML 109
 Db 1832 TW-RTNFY--NDVFTWSADAPGILYTVQHSQNNPFTAVLSQYAGAGGQFPRIV 1888
 QY 110 AGNAFTAGKIIFAAAPPVFPVENLSPS-QITMFPHVIIIDVRLTEPVLPLMPDVRSTLPHF 168
 Db 1889 AGSGVFGGLRVAVIPPGI---EIGGLEVRQFPHVIDARSLEPVTITPDLRPNMYH- 1944
 QY 169 NQKDEPKM--RLVAMLYTPLRSGDDVFTVSCRILTRPSPEPFTYLVPP---TVESK 223
 Db 1945 -PTGDDGLVPTLVLSVYNNL-INPFGGSTSAIQVTETRPSSDFEFVMIRAPSSKTVDSI 2002
 QY 224 TRP--FTLPLVLIGELNSRFPISIDEMVTSNESIV-VOPQNGRVT-----LDGBLL 273
 Db 2003 SPAGLLTTPLT-GVGNDRW-----NGQIVGLQPVPGFSTCNRHWNLSGTY 2050
 QY 274 GTTOLQACNICIRGKVTGQVPSEQHMNLEITNLNGTQFDPTDDVPAPLGVPDF----- 328
 Db 2051 GMSPPFGIDHRRGSASYSGSNATNLQFMYANAGSAIDNPISQV-APDGPDMSPVFP 2109
 QY 329 -----AGEV-FGVLSQRNGESNPAN-RAHDAVVAT-----YSDK 361
 Db 2110 NRGPIPAAGWVGAIWNSGAPNVTTVQAVELGATGACPNLQPTNTTSGAQTVAISI 2169
 QY 362 YTPKGLGIVQ-----IGTWNINDVENQPT-----KFTPIGLN 392
 Db 2170 YAVVTGAQNPAGLFWNASGIIITPNAQRITVTPQDRIVTTPGTAAAPVGNTPIMPA 2229
 QY 393 EV-----ANGHFEQWTLPRYSGALTNNMLAPAVAPLFGGERLLFFRSYVP 439
 Db 2230 SVVVRTGDVNATAGSANGTQGTGSQLP-----PVTIGLSNNYSALMPGQ-----FFVWQLT 2283
 QY 440 LKGFGNPALDCVPQEWVQHFYESAPS-----LGDVALRVYVNP 480
 Db 2284 FASGF--MEIGLSVD-----GYFAGTGASTLIDLDELIDVRVGP 2323
 RESULT 7
 A48562 coat protein - San Miguel sea lion virus (serotype 1)
 N;Alternate names: capsid protein
 C;Species: San Miguel sea lion virus
 C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
 C;Accession: A48562
 R;Neill, J.D.
 Virus Res. 24, 211-222, 1992
 A;Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lions.
 A;Reference number: A48562; MUID:92410750; PMID:1529644
 A;Accession: A48562
 A;Molecule type: genomic RNA
 A;Residues: 1-702 <NE>
 A;Cross-references: GB:M87481; NID:g334882; PIDN:AAA16217.1; PID:g334884
 A;Note: sequence extracted from NCBI backbone (NCBIN:113564, NCBI:P.113565)
 C;Superfamily: feline calicivirus coat protein
 C;Keywords: capsid protein; coat protein; glycoprotein
 F;208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted
 Query Match 11.2%; Score 318; DB 1; Length 702;
 Best Local Similarity 25.8%; Pred. No. 2.4e-15;
 Matches 153; Conservative 74; Mismatches 199; Indels 168; Gaps 33;
 QY 8 ATPSDNGAAGLVPSNNEMALEPVVGAS-----LAAPVTGQNIID-PWIRNFVQ---A 59
 Db 151 AESDGGGADIIVTEEGTVVQQPVPAQSALTTLAAAGTGKT--VDCEW--TTFFSYHTA 206
 QY 60 PNGETVSPRNSPGILLVNLGLPELNPVLAHLARMYNGYAGGMEVQVMLAGNAFTAGKI 119
 Db 207 VNWSTT-----EAQKILFRSALSPELNPLYRHSLSYSTWSGGIDVRFTVSGSVFGGL 262

QY 120 IFAAAPPVE-FVENLSPSOITMFPHVIIIDVRLTEPVLPLMPDVRSTLPHFNQKDEPKMRL 178
 Db 263 AALIVPPGTEPVE--SPTML-QYPHVLDPARQTEPVIPTIDIRKTLVH-SMDDTDITRL 318
 QY 179 VAMLYTPL-RNSGDDVFTVSCRILTRPSPEPFTYLVPPVVEKTKPFTLPLVLTLGL 237
 Db 319 VIMVYNELINPYEQSEPKSCSITVETRPSSDFTFSLKPPG-----SLKHGSI 368
 QY 238 SNRFPPLSIDEMVTSNESIVVQPPQNGR-----VTLDGELLGTTQLQACNICIRK 289
 Db 369 -----PSDLI---PRNSRHWMGNRWSTIDGFFVQVPRVQSNRHPDFDST 410
 QY 290 VTG-----QVPSRQHWMNLEITNLNGTQFDPTDD---VPA-PLGVDP 327
 Db 411 TTGWSPTPIPIEV---TLEKLDRGQYFKAVTDTEKSLVPLGPDGWPDTTITPAMTASNG 467
 QY 328 -----FAG-EVFGVLSQR-----NRGESNPANRAHDAVVATYSOKY-- 362
 Db 468 NYDYTVAEYRITNNGTHFKGYINGNLTKVKGSDNLGETQOTSRTLFLASVGNKYKQNTI 527
 QY 363 --TPKGLVVOIGTWNINDVENQPTKFT-----PIGLN-----EVANGH 398
 Db 528 NPTHKITSNLSLVYDANNVSAATAKTTTWHSTMSHLGYVLVDES PVGSDSTKVVRIA--- 584
 QY 399 RFEQWTLPRYSGALTNNMLAPAVAPLFGGERLLFFRSYVPLKGGFGNPAIDC-----S 452
 Db 585 ----TLPE--AFTNGN-----FP-----VFTNKIQI-GHFDRAHTKCFNSQVLM 623
 QY 453 VPQEWVQHFYESAPS LGDVALRVYVNPDTGRVLFEAKLHKGGLT VTSSTSGP 506
 Db 624 TSQKLAENHY-----TLPPDSLLVYRITDAASSWFDLGINHDFSYV-GISTIP 671
 RESULT 8
 S64740 genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)
 N;Contains: VP60 protein
 C;Species: rabbit hemorrhagic disease virus
 A;Variety: isolate AST/89
 C;Date: 12-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Jun-2000
 C;Accession: S64740; S46944; S49018; S65012
 R;Casais, R.; Martin-Alonso, J.; Boga, J.; Parra, F.
 submitted to the EMBL Data Library, May 1995
 A;Description: Genomic organization of rabbit hemorrhagic disease virus determined by d.
 A;Reference number: S64740
 A;Accession: S64740
 A;Molecule type: genomic RNA
 A;Residues: 1-2344 <CAS>
 A;Cross-references: EMBL:Z49271; NID:g1182032; PIDN:CAA89265.1; PID:g1150552
 A;Experimental source: isolate AST/89
 R;Boga, J.; Casais, R.; Martin, M.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra, F.
 submitted to the EMBL Data Library, July 1993
 A;Description: Molecular cloning, sequence and expression of the capsid protein gene fro
 A;Reference number: S46944
 A;Accession: S46944
 A;Molecule type: Genomic RNA
 A;Residues: 1650-2344 <BOG>
 A;Cross-references: EMBL:Z24757; NID:g515622; PIDN:CAA80881.1; PID:g515623
 A;Experimental source: isolate AST/89
 R;Parra, F.; Boga, J.A.; Martin, M.S.; Casais, R.
 Virus Res. 27, 219-228, 1993
 A;Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus supp
 A;Reference number: S49018; MUID:93255896; PMID:8488721
 A;Accession: S49018
 A;Molecule type: Genomic RNA
 A;Residues: 1650-1796 <PAR>
 A;Cross-references: EMBL:Z24757
 A;Experimental source: isolate AST/89
 A;Accession: S65012
 A;Molecule type: protein
 A;Residues: 1767-1779;1875-1877, 'X',1940-1941 <PAW>
 C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase

C:Keywords: blocked amino end; polyprotein

Query Match 11.1%; Score 316; DB 2; Length 2344;
Best Local Similarity 25.7%; Pred. No. 2e-14;
Matches 152; Conservative 80; Mismatches 212; Indels 148; Gaps 30;

QY 3 MASNDATPSDGAAGL-----VPESNNEAM-----ALEPVV-----GASLAAPVTG---- 43
DB 1766 MEGKARTAPGGAAGATTASVPGTTDGMDFGWATTSVVAENSSASIAATAGTGGPPQ 1825

QY 44 QTNIDIPWIRNFVQAPNCEFTVSPRNSPGEILVNLGPELNPYLALHARMYNGYAGGM 103
DB 1826 QVDDQETW-RTFYF--NDVFTWSVADAFGSLITVQHSFQNNPFTAVLSQMYAGWAGM 1882

QY 104 EYQVMLAGNAFTAGKIIFAAVPPYPEVENLSPS-QITMFPFHVIIIDVRLTEPVLMPDVR 162
DB 1883 QPRFIVAGIVGEGRLVAAPVPGI---BIGPGLVRFPHVVIDARSLEPVTITMPDLR 1939

QY 163 STLPHFNOKDEPKM--RLVAMLYTPLRNSGGDDVFTVSCRILTRPSPEFDTYLVPP-- 218
DB 1940 PNNYH--PTGDPGLVPTLVLSVNNL-INPFGGSTSAIQVTVETRPSEDFEFVIRAPSS 1996

QY 219 -TVESKTKP--FTLPVLTTLGELNSRPLSIDEMWTSNIESIV-VQPNQGRVT----- 267
DB 1997 KTVDSISPAGLITTPVLT-GVGNDRW-----NGQIVGLQVPFGGFTCNRHN 2044

QY 268 LDGELLGTTQLOACNICISIRKGVQVPSQEQHMNLEITNLNGTQPDPTDDVPAPLGVDP 327
DB 2045 LNSTYGVSSPRFGDIGHRRGASYPGNATVLPQWYANAGSAIDNPISQV-APDGFDP 2103

QY 328 F-----AGEV-FGVLSQNRGESN-----PAN----- 348
DB 2104 MSFVPFNGGIPAAAGVGFAGIWNSSGAPNVTTQAYELGFATGAPGNLQPTTTSQSQ 2163

QY 349 ---RAHDAVVATYSK---YTPKGLGVQIGTMTNDVENQPTKF-----TPIGLN 392
DB 2164 TVAKSIYAVVTGTAQNPAGLFWAGSVISTPSSANAITYTPQDRIVTPGTAAAPVGN 2223

QY 393 E-----VANGHRFEQWTLFRYSGALTLMNMLAPAVAPLFPGERLLF 433
DB 2224 TPIMPASVVRRTGDVNTAGSANGTYGTGSQPL---PVTIGLSLNNYSSALMPGQ---F 2277

QY 434 PRSYVPLKGGFNPAIDCSVPOEWQHFQESAPS-----LGDVALRVYVNP 480
DB 2278 FVWQLTFASGP--MEIGLSVD---GYFYAGTGCASTLIDLTELIDVVRPVG 2323

RESULT 9
A:Accession: A53982
A:Title: European brown hare syndrome virus
C:Species: European brown hare syndrome virus
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 28-Jul-2000
C:Accession: A53982
R:Wirthlich, C.; Meyers, G.; Ohlinger, V.F.; Capucci, L.; Bekens, U.; Haas, B.; Thiel, H.
J. Virol. 68, 5164-5173, 1994
A:Title: European brown hare syndrome virus: relationship to rabbit hemorrhagic disease
A:Reference number: A53982; MUID:94309183; PMID:7518531
A:Accession: A53982
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-576 <WR>
A:Cross-references: GB:U09199
C:Superfamily: human calicivirus capsid protein

Query Match 11.0%; Score 313; DB 2; Length 576;
Best Local Similarity 26.6%; Pred. No. 4.1e-15;
Matches 111; Conservative 67; Mismatches 148; Indels 92; Gaps 21;

QY 3 MASNDATPSDGAAGLVPESNNEAMALEPVVVGASLAAPVTGQTNIDIPWIRNFVQAPNG 62
DB 31 VASTDVTADNVAS-----VATAGTGGP-PQASQESW-RVNFYF--ND 72

QY 63 EFTVSPRNSPGEILVNLGPELNPYLALHARMYNGYAGGMEVQVMLAGNAFTAGKIIFA 122

DB 73 VFTSVTDAAGSILYSVQHSQNNPFTQVLSQMYAGWAGMQFRFIVAGSGIFGGRVCA 132
QY 123 AVPPYFPVENLSPS-QITMFPFHVIIIDVRLTEPVLMPDVRSTLPHFNOKDEPKM--RLV 179
DB 133 ITPPGIQIQ---PGLVRFPHVVIDARSLEPVTITMPDLRPEMYH--PTGDPGLVPTLV 187

QY 180 AMLYTPLRNSGGDDVFTVSCRILTRPSPEFDTYLVPP---TVESKTKP--TLPVLT 234
DB 188 VSVNNL-INPFGGSTSAIQVTVETRPSEDFEVLIRAPSSKTVDVSNPSMLLTTPVLT- 245

QY 235 GELNSRPLSIDEMWTSNIESIVQV-----QNGRVTLDGELLGTTQLOACNICSR 287
DB 246 GAGSDNRWAGPI-----VGLQVPFGGFTSNRHNMGSTYGMSSPRFPDDIDHPS 295

QY 288 GKV---TGQVPSQEQHMNLEITNLNGTQPDPTDDVPAPLGVDP-----PAGEV----- 332
DB 296 GNVSYPTGSATNTIETW---YANAGTATNPISNI-APDGFDMGAIPSGGTITPAGMV 351

QY 333 -FGVLSQNRGESNPNRAHDAVVATYSKYPKGLGVQIGTMTNDVENQPTKFTPI 389
DB 352 GFGQVWNASG-----TPYVGTQV---AYELGFANGAPSSIRPV 387

RESULT 10
VCWVFF
coat protein - feline calicivirus (strain CFI/68 FIV)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Aug-1999
C:Accession: A40507; B40507; T09246
R:Neill, J.D.; Reardon, I.M.; Heinrikson, R.L.
J. Virol. 65, 5440-5447, 1991
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline calicivirus
A:Reference number: A40507; MUID:91374597; PMID:1716692
A:Accession: A40507
A:Molecule type: genomic RNA
A:Residues: 1-668 <NE1>
A:Cross-references: GB:M32819; NID:G323874; PIDN:AAA42925.1; PID:G323875
A:Accession: B40507
A:Molecule type: protein
A:Residues: 373-379; 403-419; 481-489; 560-566 <NE2>
R:Neill, J.D.
submitted to the EMBL Data Library, April 1998
A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.
A:Reference number: Z16626
A:Accession: T09246
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: genomic RNA
A:Residues: 1-668 <NE3>
A:Cross-references: EMBL:U13992; NID:G3056875; PIDN:AAC13993.1; PID:G537256
A:Experimental source: strain CFI/68 FIV
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,301,304,439,459,615/Binding site: carbohydrate (Aen) (covalent) #status predicted

Query Match 10.3%; Score 292.5; DB 1; Length 668;
Best Local Similarity 26.0%; Pred. No. 1.6e-13;
Matches 113; Conservative 53; Mismatches 143; Indels 125; Gaps 21;

QY 11 SNDGAAGLVPESNN--EAMALEPVVVGASLAAPVTGQTNIDIPW-IRTNFVQAPNGEFTVS 67
DB 125 ADDGSI-TTPEQGTWVGVIAPNAQMTAADMATGKSDVSEWAEFFSFHTSVNN---S 179

QY 68 PRNSPGEILVNLGPELNPYLALHARMYNGYAGGMEVQVMLAGNAFTAGKIIFAAPVPPY 127
DB 180 TSETQGTGLFKQSLGPLNPYLTLAKLYVAVSGSDVRFISGSGVFGGKLAALVWPPG 239

QY 128 F-PVENLSPSQITMFPFHVIIIDVRLTEPVLMPDVRSTLPHFNOKDEPKMLVAMLYTPL 186
DB 240 IDPQVSTMLQ---YPHVLFDARQEPVIFSIDPLRSTLYHL-MSDITDTLSIVIMVNDL 295

QY 187 -----RNSGSGDDVFTVSCRILTRPSPEFDTYLVLPPTVESKTKPFTLPVLTGELNS 240

Db 423 TTIAEDLTTPAGDYAITSGNGNDITTTGSEYDSTEVK---NNTNFRGMYICGLQRAWGDK 479

QY 345 NPNANRAHDAVAVATYSDKTPK--LGLVQIGTWTNDV-ENQTKFTPI-----GLNEVA 395

Db 480 KISNTAFITTAKEGKIRPSNTIDMTKIAVQDTHVGEVQTSDDALALLGVGTGGEQA 539

QY 396 NGRHREQWTLPRYSALTNNMLAPAVAPLFFGERLLFRSVVPLKGGFNPADICSVPO 455

Db 540 IGSDRDR-----VVRISVLEVGAR-CGNHPIFYKNSIKL--GVVIRSIDV---- 582

QY 456 EWWQFYQBSAPSLGDAVLRV-VNPDTRVL-----FEAKLHKGGLFUTVSSTSG 505

Db 583 -----FNSQLHTSRLSLNHYLLSPDFAVVRIDNSGWFIDIGIDSEGFVGISSIG 637

QY 506 PVVVPANGYF 515

Db 638 KLEYPLSASY 647

RESULT 13

VCWVFC

coat protein - feline calicivirus (strain Japanese F4)

N:Alternate names: capsid protein

C:Species: feline calicivirus

C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000

C:Accession: B40481

R:Tohya, Y.; Taniguchi, Y.; Takahashi, E.; Utagawa, E.; Takeda, N.; Miyamura, K.; Yamazaki

Virology 183, 810-814, 1991

A:Title: Sequence analysis of the 3'-end of feline calicivirus genome.

A:Reference number: A40481; MUID:91306470; PMID:1853578

A:Accession: B40481

A:Molecule type: genomic RNA

A:Residues: 1-668 <TOH>

A:CROSS-references: GB:D90357; NID:g221264; PIDN:BAAL4371.1; PID:g221266

C:Superfamily: feline calicivirus coat protein

C:Keywords: capsid protein; coat protein; glycoprotein

F:177,301,304,399,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.9%; Score 280.5; DB 1; Length 668;

Best Local Similarity 32.0%; Pred. No. 1.3e-12;

Matches 78; Conservative 36; Mismatches 93; Indels 37; Gaps 9;

QY 9 TPNNGAAGLVPESSNEAMALEPVVGSALAAPVTGQTNIIDPW-IRTNFVQAPNGEFTVS 67

Db 132 TPEQGTAVG-----GVIAEPQAQMSTADMASGKSVDSWEAFPSFHTSVNW---S 179

QY 68 PRNSPGEILVNLELGPENLPYLAHLARMYNGYAGMEVQVLMAGNAFTAGKIIFAAVPPY 127

Db 180 TSETQKILFKQSLGPLNLPYLAHLARMYNGYAGMEVQVLMAGNAFTAGKIIFAAVPPY- 238

QY 128 FPVENLSPSQITMFPVHVIDVRLPMPDVRSTLPHFNQKDEPKMRLVAMLYTPI- 186

Db 239 -GVDPVQSTMLQYPHVLFARQVPEVFTIPDLKSTLYHV-MSDITDTTSLVIMYNDLI 296

QY 187 -----RNSNGGDDVFTVSCRILTRPSPEPDTFLVPPVTSKTPFTLPLVLTGLSLSNR 241

Db 297 NPVANDSNSSG-CIVTVE---TKGPDPKFHLKPPG-----SVLTHGSIIPSDL 341

QY 242 FPLS 245

Db 342 IPKS 345

RESULT 14

C48562

coat protein - San Miguel sea lion virus (serotype 4)

N:Alternate names: capsid protein

C:Species: San Miguel sea lion virus

C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999

C:Accession: C48562

R:Neill, J.D.

Virus Res. 24, 211-222, 1992

A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel

eins.

A:Reference number: A48562; MUID:92410750; PMID:1529644

A:Accession: C48562

A:Molecule type: genomic RNA

A:Residues: 1-703 <NEI>

A:CROSS-references: GB:M87482; NID:g334886; PIDN:AAAL6220.1; PID:g334888

A>Note: sequence extracted from NCBI backbone (NCBIP:113567)

C:Superfamily: feline calicivirus coat protein

C:Keywords: capsid protein; coat protein; glycoprotein

F:89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.6%; Score 274; DB 1; Length 703;

Best Local Similarity 28.4%; Pred. No. 4.1e-12;

Matches 93; Conservative 54; Mismatches 142; Indels 38; Gaps 14;

QY 8 ATPNSDGAAGLVPESSNEAMALEPVVG---ASLAAPVTGQTNIIDPWIRTNFVQAPNGE 63

Db 151 AESDGPSSAEIVTEEGTIVVQOQAPAPATALATATATGK-SVEQEMW-TFF--SYHTS 206

QY 64 FTVSPRNSPGEILVNLELGPENLPYLAHLARMYNGYAGMEVQVLMAGNAFTAGKIIFAA 123

Db 207 INWSTVESQKILYVQALNPSINPYLDHIAKLYSTWSGIDVRFTVSGSGVFGKLAALL 266

QY 124 VPP-YFPVENLSPSQITMFPVHVIDVRLPMPDVRSTLPHFNQKDEPKMRLVAMLY 182

Db 267 VPPGVEPIESVSMLO---YPHVLFARQVPEVFTIPDLKSTLYHV-MSDITDTTSLVIM 322

QY 183 YTLRNSGDDVFTVSCRILTRPSPEPDTFLVPP--TVESKTKPFTLPLVLTGLSLSNR 240

Db 323 Y-----ENGVTCSITVETRPADFTFALLKPGSLIKHGSIPSLIPNSAHWMGN 377

QY 241 RPLSIDEMVTSPNSIYVQPNQGRVTLTGELLGTTQLQACNICIRGKVTQGVSEQHM 300

Db 378 RWSSTISGFSVQPR---VFQSNRHFDFDST---TTGWSPTVYVPIEIKIQKGVGNNKW 430

QY 301 WNLEITNLNGTQFDPTDDVPAPLGPVD 327

Db 431 FHV-----IDTDKALVP-GIPD 446

RESULT 15

JQ2356

capsid protein - feline calicivirus (strain KCD)

C:Species: feline calicivirus

C>Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999

C:Accession: JQ2356

R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.

J. Gen. Virol. 74, 2519-2524, 1993

A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable

A:Reference number: JQ2354; MUID:94065683; PMID:7504075

A:Accession: JQ2356

A:Molecule type: mRNA

A:Residues: 1-668 <SEA>

A:CROSS-references: GB:L09719; NID:g305107; PIDN:AAAL6487.1; PID:g305108

C:Superfamily: feline calicivirus coat protein

C:Keywords: capsid protein

Query Match 9.5%; Score 269; DB 2; Length 668;

Best Local Similarity 23.9%; Pred. No. 8.8e-12;

Matches 137; Conservative 83; Mismatches 234; Indels 120; Gaps 28;

QY 11 SNDGAAGLVPESSNN--EAMALEPVVGSALAAPVTGQTNIIDPW-IRTNFVQAPNGEFTVS 67

Db 125 ADDGSI-VTPEQGTWVGIVAEPAQMAQSTADMATKSVDSWEAFPSFHTSVNW---S 179

QY 68 PRNSPGEILVNLELGPENLPYLAHLARMYNGYAGMEVQVLMAGNAFTAGKIIFAAVPPY 127

Db 180 TSETQKILFKQSLGPLNLPYLAHLARMYNGYAGMEVQVLMAGNAFTAGKIIFAAVPPY- 238

QY 128 FPVENLSPSQITMFPVHVIDVRLPMPDVRSTLPHFNQKDEPKMRLVAMLYTPI- 186

Db 239 -GVDPVQSTMLQYPHVLFARQVPEVFTIPDLKSTLYHV-MSDITDTTSLVIMYNDLI 296

Search completed: June 1, 2004, 13:55:27
 . Job time : 12.4195 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2811	98.8	540	12	Q917X1	Q917x1 norwalk-lik
2	2806	98.6	538	12	Q91165	Q91165 human calic
3	2796	98.3	540	12	Q917M8	Q917M8 norwalk-lik
4	2789	98.0	540	12	Q91V50	Q91V50 human calic
5	2361	83.0	542	12	Q12305	Q12305 snow mounta
6	2361	83.0	542	12	Q80RD6	Q80rd6 snow mounta
7	2360	83.0	542	12	Q96877	Q96877 snow mounta
8	2351	82.6	542	12	Q913B6	Q913b6 human calic
9	2339	82.2	542	12	Q66293	Q66293 calicivirus
10	2336	81.8	548	12	Q915C7	Q915c7 human calic
11	2273.5	79.9	535	12	Q91115	Q91115 human calic
12	2270.5	78.8	535	12	Q68104	Q68104 hawaii cali
13	2269.5	79.8	535	12	Q9QMK6	Q9qmk6 chitta viru
14	2268.5	79.7	535	12	Q915C6	Q915c6 human calic
15	2268.5	79.7	535	12	Q917X7	Q917x7 norwalk-lik
16	2261.5	79.5	535	12	Q917X4	Q917x4 norwalk-lik

```

J. Infect. Dis. 181:S336-S348 (2000).
[4]
SEQUENCE FROM N.A.
STRAIN=Hu/NLJ/New Orleans/306/1994/US;
Ando T., Sato Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
EMBL; AF414422; AAL13019.1; -
DR InterPro; IPR004005; Calici.coat.
DR Pfam; PF00915; Calici.coat.1; 7C3DFBF3893896E5 CRC64;
SQ SEQUENCE 540 AA; 58866 MW; 58866 MW; F0D0B52BBC0E6024 CRC64;

Query Match 98.8%; Score 2811; DB 12; Length 540;
Best Local Similarity 98.9%; Pred. No. 3e-214;
Matches 534; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MKWASNDATPSNDGAAGLVPSNEAMALEPVGASLAAPVTGOTNIIDPWIRTNFVQAP 60
DB 1 MKWASNDATPSNDGAAGLVPSNEAMALEPVGASLAAPVTGOTNIIDPWIRTNFVQAP 60

QY 61 NGEFTVSPRNSPCEILLVNLELGPENLPYLALHARMYNGYAGGMEVQVMLAGNAFTAGKII 120
DB 61 NGEFTVSPRNSPCEILLVNLELGPENLPYLALHARMYNGYAGGMEVQVMLAGNAFTAGKII 120

QY 121 FAVAPYFPVENLSPSQITMFPFHVIIDVRLTEPVLPMDDVRSITLPHFNQKDEPKRLVA 180
DB 121 FAVAPYFPVENLSPSQITMFPFHVIIDVRLTEPVLPMDDVRSITLPHFNQKDEPKRLVA 180

QY 181 MLYTPLRNSGDDVFTVSCRILTRPSPEDFYLVPPVTSKTKPFTLPVLTGLGELNS 240
DB 181 MLYTPLRNSGDDVFTVSCRILTRPSPEDFYLVPPVTSKTKPFTLPVLTGLGELNS 240

QY 241 RFLSIDEMVTSRNESIVVQPNQGRVTLTGELLTGTTQACNICSRIGKVTGQVPSQHM 300
DB 241 RFLSIDEMVTSRNESIVVQPNQGRVTLTGELLTGTTQACNICSRIGKVTGQVPSQHM 300

QY 301 WNLITNLNGTQDPTDDVPAPLGVDPFAGEVGVLSQRNAGESNPANRAHADVAATYSD 360
DB 301 WNLITNLNGTQDPTDDVPAPLGVDPFAGEVGVLSQRNAGESNPANRAHADVAATYSD 360

QY 361 KYTPKGLGLVQIGTWNINDVENQPTKFTPIGLNEVANGHREFEQWTLPRYSGALTLNMLAP 420
DB 361 KYTPKGLGLVQIGTWNINDVENQPTKFTPIGLNEVANGHREFEQWTLPRYSGALTLNMLAP 420

QY 421 AVAPLFGGERLLFRSVPVPLKGGFGNPAIDCVSPQEWVQHFYQESAPSLGDVALVRYNP 480
DB 421 AVAPLFGGERLLFRSVPVPLKGGFGNPAIDCVSPQEWVQHFYQESAPSLGDVALVRYNP 480

QY 481 DTGRVLFKHLKGGFLTVSSTSTGVPVVPANGYFKFDSWVNOFYSLAPMGTTGNGRRRQ 540
DB 481 DTGRVLFKHLKGGFLTVSSTSTGVPVVPANGYFKFDSWVNOFYSLAPMGTTGNGRRRQ 540

RESULT 2
Q91165 PRELIMINARY; PRT; 538 AA.
ID Q91165
AC Q91165;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/MOH/99.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=168224;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLJ/MOH/99;
RA Faras T., Berke T., Reuter G., Szucs G., Matson D., Jiang X.;
RT "Molecular detection and sequence analysis of human caliciviruses from acute gastroenteritis outbreaks in Hungary.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF397156; AAK84404.1; -

J. Infect. Dis. 181:S336-S348 (2000).
[4]
SEQUENCE FROM N.A.
STRAIN=Hu/NLJ/New Orleans/306/1994/US;
Ando T., Sato Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
EMBL; AF414422; AAL13019.1; -
DR InterPro; IPR004005; Calici.coat.
DR Pfam; PF00915; Calici.coat.1; 7C3DFBF3893896E5 CRC64;
SQ SEQUENCE 540 AA; 58866 MW; 58866 MW; F0D0B52BBC0E6024 CRC64;

Query Match 98.8%; Score 2811; DB 12; Length 540;
Best Local Similarity 98.9%; Pred. No. 3e-214;
Matches 534; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MKWASNDATPSNDGAAGLVPSNEAMALEPVGASLAAPVTGOTNIIDPWIRTNFVQAP 60
DB 1 MKWASNDATPSNDGAAGLVPSNEAMALEPVGASLAAPVTGOTNIIDPWIRTNFVQAP 60

QY 61 NGEFTVSPRNSPCEILLVNLELGPENLPYLALHARMYNGYAGGMEVQVMLAGNAFTAGKII 120
DB 61 NGEFTVSPRNSPCEILLVNLELGPENLPYLALHARMYNGYAGGMEVQVMLAGNAFTAGKII 120

QY 121 FAVAPYFPVENLSPSQITMFPFHVIIDVRLTEPVLPMDDVRSITLPHFNQKDEPKRLVA 180
DB 121 FAVAPYFPVENLSPSQITMFPFHVIIDVRLTEPVLPMDDVRSITLPHFNQKDEPKRLVA 180

QY 181 MLYTPLRNSGDDVFTVSCRILTRPSPEDFYLVPPVTSKTKPFTLPVLTGLGELNS 240
DB 181 MLYTPLRNSGDDVFTVSCRILTRPSPEDFYLVPPVTSKTKPFTLPVLTGLGELNS 240

QY 241 RFLSIDEMVTSRNESIVVQPNQGRVTLTGELLTGTTQACNICSRIGKVTGQVPSQHM 300
DB 241 RFLSIDEMVTSRNESIVVQPNQGRVTLTGELLTGTTQACNICSRIGKVTGQVPSQHM 300

QY 301 WNLITNLNGTQDPTDDVPAPLGVDPFAGEVGVLSQRNAGESNPANRAHADVAATYSD 360
DB 301 WNLITNLNGTQDPTDDVPAPLGVDPFAGEVGVLSQRNAGESNPANRAHADVAATYSD 360

QY 361 KYTPKGLGLVQIGTWNINDVENQPTKFTPIGLNEVANGHREFEQWTLPRYSGALTLNMLAP 420
DB 361 KYTPKGLGLVQIGTWNINDVENQPTKFTPIGLNEVANGHREFEQWTLPRYSGALTLNMLAP 420

QY 421 AVAPLFGGERLLFRSVPVPLKGGFGNPAIDCVSPQEWVQHFYQESAPSLGDVALVRYNP 480
DB 421 AVAPLFGGERLLFRSVPVPLKGGFGNPAIDCVSPQEWVQHFYQESAPSLGDVALVRYNP 480

QY 481 DTGRVLFKHLKGGFLTVSSTSTGVPVVPANGYFKFDSWVNOFYSLAPMGTTGNGRRRQ 540
DB 481 DTGRVLFKHLKGGFLTVSSTSTGVPVVPANGYFKFDSWVNOFYSLAPMGTTGNGRRRQ 540

RESULT 2
Q91165 PRELIMINARY; PRT; 538 AA.
ID Q91165
AC Q91165;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/MOH/99.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=168224;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLJ/MOH/99;
RA Faras T., Berke T., Reuter G., Szucs G., Matson D., Jiang X.;
RT "Molecular detection and sequence analysis of human caliciviruses from acute gastroenteritis outbreaks in Hungary.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF397156; AAK84404.1; -

J. Infect. Dis. 181:S336-S348 (2000).
[4]
SEQUENCE FROM N.A.
STRAIN=Hu/NLJ/New Orleans/306/1994/US;
Ando T., Sato Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
EMBL; AF414422; AAL13019.1; -
DR InterPro; IPR004005; Calici.coat.
DR Pfam; PF00915; Calici.coat.1; 7C3DFBF3893896E5 CRC64;
SQ SEQUENCE 540 AA; 58866 MW; 58866 MW; F0D0B52BBC0E6024 CRC64;

Query Match 98.6%; Score 2806; DB 12; Length 538;
Best Local Similarity 98.9%; Pred. No. 7.5e-214;
Matches 532; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 MASNDATPSNDGAAGLVPSNEAMALEPVGASLAAPVTGOTNIIDPWIRTNFVQAPNG 62
DB 1 MASNDATPSNDGAAGLVPSNEAMALEPVGASLAAPVTGOTNIIDPWIRTNFVQAPNG 60

QY 63 EFTVSPRNSPCEILLVNLELGPENLPYLALHARMYNGYAGGMEVQVMLAGNAFTAGKII 122
DB 61 EFTVSPRNSPCEILLVNLELGPENLPYLALHARMYNGYAGGMEVQVMLAGNAFTAGKII 120

QY 123 AVPPYFPVENLSPSQITMFPFHVIIDVRLTEPVLPMDDVRSITLPHFNQKDEPKRLVAM 182
DB 121 AVPPYFPVENLSPSQITMFPFHVIIDVRLTEPVLPMDDVRSITLPHFNQKDEPKRLVAM 180

QY 183 YTPLRNSGDDVFTVSCRILTRPSPEDFYLVPPVTSKTKPFTLPVLTGLGELNSRF 242
DB 181 YTPLRNSGDDVFTVSCRILTRPSPEDFYLVPPVTSKTKPFTLPVLTGLGELNSRF 240

QY 243 PLSIDEMVTSRNESIVVQPNQGRVTLTGELLTGTTQACNICSRIGKVTGQVPSQHMWN 302
DB 241 PLSIDEMVTSRNESIVVQPNQGRVTLTGELLTGTTQACNICSRIGKVTGQVPSQHMWN 300

QY 303 LEITNLNGTQDPTDDVPAPLGVDPFAGEVGVLSQRNAGESNPANRAHADVAATYSDKY 362
DB 301 LEITNLNGTQDPTDDVPAPLGVDPFAGEVGVLSQRNAGESNPANRAHADVAATYSDKY 360

QY 363 TPKGLGLVQIGTWNINDVENQPTKFTPIGLNEVANGHREFEQWTLPRYSGALTLNMLAP 422
DB 361 TPKGLGLVQIGTWNINDVENQPTKFTPIGLNEVANGHREFEQWTLPRYSGALTLNMLAP 420

QY 423 APLFGGERLLFRSVPVPLKGGFGNPAIDCVSPQEWVQHFYQESAPSLGDVALVRYNP 482
DB 421 APLFGGERLLFRSVPVPLKGGFGNPAIDCVSPQEWVQHFYQESAPSLGDVALVRYNP 480

QY 483 GRVLFEAKLHKGGLTVSSTSTGVPVVPANGYFKFDSWVNOFYSLAPMGTTGNGRRRQ 540
DB 481 GRVLFEAKLHKGGLTVSSTSTGVPVVPANGYFKFDSWVNOFYSLAPMGTTGNGRRRQ 538

RESULT 3
Q917W8 PRELIMINARY; PRT; 540 AA.
ID Q917W8
AC Q917W8;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Capsid protein.
OS Norwalk-like virus NLV/White River/290/1994/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171855;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLJ/White River/290/1994/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify a 3-kilobase region from the RNA polymerase gene to the poly(A) tail of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577 (1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLJ/White River/290/1994/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
```

RA Seto Y., Monroe S.S., Glass R.I.;
 RT "Correlation of patient immune responses with genetically
 RT characterized small round-structured viruses involved in outbreaks of
 RT nonbacterial acute gastroenteritis in the United States, 1990 to
 RT 1995.";
 RN J. Med. Virol. 53:372-383(1997).
 RW [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Hu/NLV/White River/290/1994/US;
 RX MEDLINE=20266071; PubMed=10804147;
 RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
 RA Clegg J.C., Chamberlain J., Brown D.W.G.;
 RT "Capsid protein diversity among 'Norwalk-like' viruses.";
 RL Virus Genes 20:227-236(2000).
 RN J. Infect. Dis. 181:S336-S348(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Hu/NLV/White River/230/1994/US;
 RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF414423; AAL13022.1; --
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 SQ SEQUENCE 540 AA; 58788 MW; 9C8164BEA52F53B9 CRC64;

Query Match 98.3%; Score 2796; DB 12; Length 540;
 Best Local Similarity 98.5%; Pred. No. 4.7e-213;
 Matches 532; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
 QY 1 MKMASNDATPSNDGAAGLVPESSNNEAMALEPVGASLAAPVTGQTNIIDPWIRTFVQAP 60
 DB 1 MKMASNDATPSNDGAAGLVPESSNNEAMALEPVGASLAAPVTGQTNIIDPWIRTFVQAP 60
 QY 61 NGEFTVSPNSPGEILVNLELGPENLPYLAHLARMYNGYAGGMEVQVMLAGNAFTAGKII 120
 DB 61 NGEFTVSPNSPGEILVNLELGPENLPYLAHLARMYNGYAGGMEVQVMLAGNAFTAGKII 120
 QY 121 FAAVPPYFPVENLSPSQITMFPFHVIDVRLTLEPVLPMDDVRSSTLPHFNQKDEPKMLVA 180
 DB 121 FAAVPPYFPVENLSPSQITMFPFHVIDVRLTLEPVLPMDDVRSSTLPHFNQKDEPKMLVA 180
 QY 181 MLYTPLRSNGSDDDVFTVSCRLLTRPSPEFDTYLVPPVTSKTKPFTLPVLTGLGELS 240
 DB 181 MLYTPLRSNGSDDDVFTVSCRLLTRPSPEFDTYLVPPVTSKTKPFTLPVLTGLGELS 240
 QY 241 RPFLSIDEMVTSNPESIVVQPNQGRVTLTGELLGTTQLQACNICSIKRGVTCQVSEQHM 300
 DB 241 RPFLSIDEMVTSNPESIVVQPNQGRVTLTGELLGTTQLQACNICSIKRGVTCQVSEQHM 300
 QY 301 WNLEITNLNGTQFDDTDDVPAPLGVDPDFAGEVFGVLSQRNGESNPANRAHDAVVATYSD 360
 DB 301 WNLEITNLNGTQFDDTDDVPAPLGVDPDFAGEVFGVLSQRNGESNPANRAHDAVVATYSD 360
 QY 361 KYTPKGLGVOIGTWNNDVENQPTKFTPIGLNEVANGHRFEQWTLPRYSGALTANNLAP 420
 DB 361 KYTPKGLGVOIGTWNNDVENQPTKFTPIGLNEVANGHRFEQWTLPRYSGALTANNLAP 420
 QY 421 AVAPLPPGERLLFFRSYVPLKGGFNGNPAIDCSVPQEWQHFYQESAPSLGDVALVRYNVP 480
 DB 421 AVAPLPPGERLLFFRSYVPLKGGFNGNPAIDCSVPQEWQHFYQESAPSLGDVALVRYNVP 480
 QY 481 DTGRVLFEAKLHKGFLTVSSTSTGPPVVVPANGYKFKDSWVNFQYSLAPMGTCNGRRRVQ 540
 DB 481 DTGRVLFEAKLHKGFLTVSSTSTGPPVVVPANGYKFKDSWVNFQYSLAPMGTCNGRRRVQ 540

RESULT 4
 Q9IV50
 ID Q9IV50 PRELIMINARY; PRT; 540 AA.
 AC Q9IV50;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Capsid protein.

OS Human calicivirus HU/NLV/Hillingdon/90/UK.
 OC Viruses; sRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 RN NCBI_TaxID=122911;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Hu/NLV/Hillingdon/90/UK;
 RX MEDLINE=20404883; PubMed=10949950;
 RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
 RA Clegg J.C., Chamberlain J., Brown D.W.G.;
 RT "Capsid protein diversity among 'Norwalk-like' viruses.";
 RL Virus Genes 20:227-236(2000).
 RN EMBL; AJ277607; CAB89088.1; --
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 SQ SEQUENCE 540 AA; 59043 MW; B224257E68B0DDCD CRC64;

Query Match 98.0%; Score 2789; DB 12; Length 540;
 Best Local Similarity 98.1%; Pred. No. 1.7e-212;
 Matches 530; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 QY 1 MKMASNDATPSNDGAAGLVPESSNNEAMALEPVGASLAAPVTGQTNIIDPWIRTFVQAP 60
 DB 1 MKMASNDATPSNDGAAGLVPESSNNEAMALEPVGASLAAPVTGQTNIIDPWIRTFVQAP 60
 QY 61 NGEFTVSPNSPGEILVNLELGPENLPYLAHLARMYNGYAGGMEVQVMLAGNAFTAGKII 120
 DB 61 NGEFTVSPNSPGEILVNLELGPENLPYLAHLARMYNGYAGGMEVQVMLAGNAFTAGKII 120
 QY 121 FAAVPPYFPVENLSPSQITMFPFHVIDVRLTLEPVLPMDDVRSSTLPHFNQKDEPKMLVA 180
 DB 121 FAAVPPYFPVENLSPSQITMFPFHVIDVRLTLEPVLPMDDVRSSTLPHFNQKDEPKMLVA 180
 QY 181 MLYTPLRSNGSDDDVFTVSCRLLTRPSPEFDTYLVPPVTSKTKPFTLPVLTGLGELS 240
 DB 181 MLYTPLRSNGSDDDVFTVSCRLLTRPSPEFDTYLVPPVTSKTKPFTLPVLTGLGELS 240
 QY 241 RPFLSIDEMVTSNPESIVVQPNQGRVTLTGELLGTTQLQACNICSIKRGVTCQVSEQHM 300
 DB 241 RPFLSIDEMVTSNPESIVVQPNQGRVTLTGELLGTTQLQACNICSIKRGVTCQVSEQHM 300
 QY 301 WNLEITNLNGTQFDDTDDVPAPLGVDPDFAGEVFGVLSQRNGESNPANRAHDAVVATYSD 360
 DB 301 WNLEITNLNGTQFDDTDDVPAPLGVDPDFAGEVFGVLSQRNGESNPANRAHDAVVATYSD 360
 QY 361 KYTPKGLGVOIGTWNNDVENQPTKFTPIGLNEVANGHRFEQWTLPRYSGALTANNLAP 420
 DB 361 KYTPKGLGVOIGTWNNDVENQPTKFTPIGLNEVANGHRFEQWTLPRYSGALTANNLAP 420
 QY 421 AVAPLPPGERLLFFRSYVPLKGGFNGNPAIDCSVPQEWQHFYQESAPSLGDVALVRYNVP 480
 DB 421 AVAPLPPGERLLFFRSYVPLKGGFNGNPAIDCSVPQEWQHFYQESAPSLGDVALVRYNVP 480
 QY 481 DTGRVLFEAKLHKGFLTVSSTSTGPPVVVPANGYKFKDSWVNFQYSLAPMGTCNGRRRVQ 540
 DB 481 DTGRVLFEAKLHKGFLTVSSTSTGPPVVVPANGYKFKDSWVNFQYSLAPMGTCNGRRRVQ 540

RESULT 5
 O12305
 ID O12305 PRELIMINARY; PRT; 542 AA.
 AC O12305;
 DT 01-JUL-1997 (TReMBLrel. 04, Created)
 DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE SMA capsid protein (fragment).
 OS Snow Mountain virus.
 OC Viruses; sRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 OX NCBI_TaxID=52276;
 RN [1]
 RP SEQUENCE FROM N.A.

```

RC STRAIN=Snow Mountain strain;
RX MEDLINE=97412813; PubMed=9267456;
RA Hardy M.E., Kramer S.F., Treanor J.J., Estes M.K.;
RT "Human calicivirus genotype II capsid sequence diversity revealed by
RT analyses of the prototype Snow Mountain agent.";
RL Arch. Virol. 142:1469-1479 (1997).
RE EMBL; U75682; AAB61685.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR NON_TER 542 542
SQ SEQUENCE 542 AA; 59275 MW; E0C10C4B84DAFDB1 CRC64;

Query Match 83.0%; Score 2361; DB 12; Length 542;
Best Local Similarity 79.3%; Pred. No. 1.5e-178; Indels 8; Gaps 3;
Matches 432; Conservative 57; Mismatches 48;

QY 1 MKMASNDATPSNDGAAGLVPESSNEAMALEPVGASLAAPVTGQTNIIIDPWIRTFVQAP 60
DB 1 MKMASNDAAAPSTDGAAGLVPESSNEAMALEPVGASLAAPVTGQTNIIIDPWIRTFVQAP 60
QY 61 NGEFTVSPRNSPCEILVNLELGPENLPHYLALARMYNGYAGGMEVQVMLAGNAFTAGKII 120
DB 61 NGEFTVSPRNPAGEVLNLELGPENLPHYLALARMYNGYAGGMEVQVMLAGNAFTAGKII 120
QY 121 FAAPVPFVVENLSPQITMFPFHVIIIDVRLTLEPVLPMDFVSTLPHFNQKDEPKMRLVA 180
DB 121 FAAPVPFVVENLSPQITMFPFHVIIIDVRLTLEPVLPMDFVSTLPHFNQKDEPKMRLVA 180
QY 181 MLYTPLRNSGGDDVFTVSCRIILTRSPDFDFTYLPVPTVESKTKDFTLPLVLTGLSLS 240
DB 181 MLYTPLRNSGGDDVFTVSCRIILTRSPDFDFTYLPVPTVESKTKDFTLPLVLTGLSLS 240
QY 241 RFLPSIDEMWTSNESPESIVQPNQGRVTLDELLGTTQLQACNICSIRGKVTGVPSEQHM 300
DB 241 RFPVSIIDQMTSPNEVSVQCNQGRCTLDGELGTTQLQVSGICAFKGEVTAHLQNDHL 300
QY 301 WNLKITNLNGTQFDPTDDVPAPLGVDPFAGEVGVLSQRNR----GESNPANRAHADVA 356
DB 301 YNITITNLNGSPDFSEDIAPLGVDPFQGRVFGVITQRDKQNAAGSQPANRGHDVAVP 360
QY 357 TYSKYTPKGLVQIGTWTNNDVE--NPTKFTPIGLNEVANGHRFEQWTLPRYSGLATLN 415
DB 361 TYTAQYTPKGLVQIGTWTNNDVE--NPTKFTPIGLNEVANGHRFEQWTLPRYSGLATLN 417
QY 416 MNLAPAVAPLPGERLLFFRSYVPLKGGFNGNPAIDCSVPQEWVQHFYQESAPSLGDVALV 475
DB 418 TNLAPSVAPVPFGERLLFFRSYVPLKGGFNGNPAIDCSVPQEWVQHFYQESAPSLGDVALV 477
QY 476 RYVNPDTGRVLFEAKLHKGGLTYSSTSTGTVVVPANGYKFDQSWVQNFYSLAPMGTCNG 535
DB 478 RYINPDTGRALFEAKLHKGGLTYSSTSTGTVVVPANGYKFDQSWVQNFYSLAPMGTCNG 537
QY 536 RRRVQ 540
DB 538 RRRIQ 542

RESULT 7
Q96877 PRELIMINARY; PRT; 542 AA.
AC Q96877;
DT 01-FEB-1997 (Tremblrel. 02, Created)
DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Capsid protein.
OS Snow Mountain virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=52276;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Snow Mountain strain;
RX MEDLINE=98014353; PubMed=9354262;
RA King A.D., Green K.Y.;
RT "Sequence analysis of the gene encoding the capsid protein of the Snow
RT Mountain human calicivirus.";

```

RL Virus Genes 15:5-7(1997).
DR EMBL; U70059; AAB16915.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 542 AA; 59201 MW; E7A22C488ABFDB1 CRC64;

Query Match 83.0%; Score 2360; DB 12; Length 542;
Best Local Similarity 79.1%; Pred. No. 1.8e-178;
Matches 433; Conservative 56; Mismatches 48; Indels 8; Gaps 3;

QY 1 MKMASNDATPSNDGAGLVPSNNEAMALEPVVGSALAAPVTGQTNIIIDPWIRTFVQAP 60
DB 1 MKMASNDAASTDDGAGLVPSNNEAMALEPVVGAALAAPVTGQTNIIIDPWIRTFVQAP 60

QY 61 NGEFTVSPRNSPGEITLVNLELGPENLPYLALHARMYNGYAGMEVQVMLAGNAFTAGKII 120
DB 61 NGEFTVSPRNPAGEVLLNLELGPENLPYLALHARMYNGYAGMEVQVMLAGNAFTAGKLV 120

QY 121 FFAVPPYFVENLSQITMFPFHVIDVRLPVLVLPDPVSTLTFHFNQKDEPKRLVA 180
DB 121 FFAVPPHFFVENLSQITMFPFHVIDVRLPVLVLPDPVSTLTFHFNQKDEPKRRIVA 180

QY 181 MLYTLRSNGSGDDVFTVSCRILTRPSPEFDTYLPVPTVESKTKPFTLPVLTGLSLS 240
DB 181 MLYTLRSNGSGDDVFTVSCRILTRPSPEFDTYLPVPTVESKTKPFTLPVLTGLSLS 240

QY 241 RPPLSIDENMTSPNESIVVQPNQGRVTLTGELGTTQLQACNICSRIGKVTGQVPSQHM 300
DB 241 RPVSIDQMTSPNEVISVQCNQGRCTLDGELGTTQLQVSGICASKGEVTAHLQNDHL 300

QY 301 WNLEITNLNGTQFDPTDDVPAPLGVDPDFAGEVFGVLSQRNR---GESNPANRAHDVVA 356
DB 301 YNITITNLNGSPFDSEDIAPLGVDPDFQGRVFGVITQRDKQNAAGSQSPANKRGHDVVP 360

QY 357 TYSKXTPKLGLVQIGTWTNDVE-NQPTKFTPIGLNEVANGHRFEQWTLPRYSALTIN 415
DB 361 TYTAQYTPKLGQVIGTWTQDLDLKNQPVKFTFVGLNDT---EHFNQWVVPYAGALNIN 417

QY 416 MNLAPAVAPLPGERLLPFRSVVPLKGGFGNPAIDCSVPQEWQHFYQBSAPSLGDVALV 475
DB 418 TNLAPSAVAPVFGERLLPFRSYLPLKGGYGNPAIDCLLPQEWQHFYQBAASMSSEVALV 477

QY 476 RYVNPDTGRVLFEAKLHKGGFTVSTSTGPPVVPVANGYFKFDSWMNQFYSLAPMGNG 535
DB 478 RYINPDTGRALFEAKLHRAGFTVSSNTSAPVVPVANGYFRFDSWMNQFYSLAPMGNG 537

QY 536 RRRVQ 540
DB 538 RRRVQ 542

RESULT 8
Q913B6 PRELIMINARY; PRT; 542 AA.
AC Q913B6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NUV/Chesterfield/434/1997/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=173724;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/Chesterfield/434/1997/US;
RX MEDLINE=99246332; PubMed=10228052;
RA Fankhauser R.L., Noel J.S., Monroe S.S., Ando T., Glass R.I.;
RT "Molecular epidemiology of 'Norwalk-like viruses' in outbreaks of
RT gastroenteritis in the United States.";
RL J. Infect. Dis. 178:1571-1578(1998).
RN [2]

RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=NLV/Chesterfield/434/1997/US;
RX MEDLINE=99246332; PubMed=10228052;
RA Noel J.S., Fankhauser R.L., Ando T., Monroe S.S., Glass R.I.;
RT "Identification of a distinct common strain of 'Norwalk-like viruses'
RT having a global distribution.";
RL J. Infect. Dis. 179:1334-1344(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/Chesterfield/434/1997/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses'.";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/Chesterfield/434/1997/US;
RA Noel J.S., Fankhauser R.L., Monroe S.S., Glass R.I., Ando T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY054300; AAL15548.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 542 AA; 59233 MW; 15652994B5631E31 CRC64;

Query Match 82.6%; Score 2351; DB 12; Length 542;
Best Local Similarity 79.1%; Pred. No. 9.1e-178;
Matches 431; Conservative 56; Mismatches 50; Indels 8; Gaps 3;

QY 1 MKMASNDATPSNDGAGLVPSNNEAMALEPVVGSALAAPVTGQTNIIIDPWIRTFVQAP 60
DB 1 MKMASNDAASTDDGAGLVPSNNEAMALEPVVGAALAAPVTGQTNIIIDPWIRTFVQAP 60

QY 61 NGEFTVSPRNSPGEITLVNLELGPENLPYLALHARMYNGYAGMEVQVMLAGNAFTAGKII 120
DB 61 NGEFTVSPRNPAGEVLLNLELGPENLPYLALHARMYNGYAGMEVQVMLAGNAFTAGKLV 120

QY 121 FFAVPPYFVENLSQITMFPFHVIDVRLPVLVLPDPVSTLTFHFNQKDEPKRLVA 180
DB 121 FFAVPPHFFVENLSQITMFPFHVIDVRLPVLVLPDPVSTLTFHFNQKDEPKRRIVA 180

QY 181 MLYTLRSNGSGDDVFTVSCRILTRPSPEFDTYLPVPTVESKTKPFTLPVLTGLSLS 240
DB 181 MLYTLRSNGSGDDVFTVSCRILTRPSPEFDTYLPVPTVESKTKPFTLPVLTGLSLS 240

QY 241 RPPLSIDENMTSPNESIVVQPNQGRVTLTGELGTTQLQACNICSRIGKVTGQVPSQHM 300
DB 241 RPVSIDQMTSPNEVISVQCNQGRCTLDGELGTTQLQVSGICASKGEVTAHLQNDHL 300

QY 301 WNLEITNLNGTQFDPTDDVPAPLGVDPDFAGEVFGVLSQRNR---GESNPANRAHDVVA 356
DB 301 YNITITNLNGSPFDSEDIAPLGVDPDFQGRVFGVITQRDKQNAAGSQSPANKRGHDVVP 360

QY 357 TYSKXTPKLGLVQIGTWTNDVE-NQPTKFTPIGLNEVANGHRFEQWTLPRYSALTIN 415
DB 361 TYTAQYTPKLGQVIGTWTQDLDLKNQPVKFTFVGLNDT---EHFNQWVVPYAGALNIN 417

QY 416 MNLAPAVAPLPGERLLPFRSVVPLKGGFGNPAIDCSVPQEWQHFYQBSAPSLGDVALV 475
DB 418 TNLAPSAVAPVFGERLLPFRSYLPLKGGYGNPAIDCLLPQEWQHFYQBAASMSSEVALV 477

QY 476 RYVNPDTGRVLFEAKLHKGGFTVSTSTGPPVVPVANGYFKFDSWMNQFYSLAPMGNG 535
DB 478 RYINPDTGRALFEAKLHRAGFTVSSNTSAPVVPVANGYFRFDSWMNQFYSLAPMGNG 537

QY 536 RRRVQ 540
DB 538 RRRVQ 542

RESULT 9
Q66293
ID Q66293 PRELIMINARY; PRT; 542 AA.

Q66293;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Capsid protein.
 OS Human calicivirus NLV/Erfurt/546/00/DB.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Calicivirus.
 OX NCBI_TaxID=11975;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Melksham Virus;
 RX MEDLINE=96136658; PubMed=8533462;
 RA Green S.M., Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;
 RT "Capsid diversity in small round-structured viruses : molecular
 RT characterization of an antigenically distinct human enteric
 RT calicivirus";
 RL Virus Res. 37:271-283(1995).
 DR EMBL; X81879; CAA57462.1; -;
 DR FIP; S60616; S60616.
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_Cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 SQ SEQUENCE 542 AA; 59147 MW; 0A0AA91343E60341 CRC64;
 Query Match 82.2%; Score 2339; DB 12; Length 542;
 Best Local Similarity 78.9%; Pred. No. 8.1e-177;
 Matches 430; Conservative 56; Mismatches 51; Indels 8; Gaps 3;
 QY 1 MKMASNDATPSNDGAAGLVPESSNEAMALEPVPVVGASLAAPVTGQTNIIIDPWIRTNFVQAP 60
 DB 1 MKMASNDAASTDGAAGLVPESSNEAMALEPVPVVGASLAAPVTGQTNIIIDPWIRTNFVQAP 60
 QY 61 NGEFTVSPRNSPGEIILVNLELGPENLPYLAHARMYNGVAGMEVQVMLAGNAFTAGKII 120
 DB 61 NGEFTVSPRNPAGEVLLNLELGPENLPYLAHARMYNGVAGMEVQVMLAGNAFTAGKII 120
 QY 121 FAAPVPFPVENLSPSQITMFPFHVIIIDVRLPEVLLPMPDVRSTLPHNQDEPKRLVA 180
 DB 121 FAAPVPFPVENLSPSQITMFPFHVIIIDVRLPEVLLPMPDVRSTLPHNQDEPKRLVA 180
 QY 181 MLYTPLRSNGSGDDVFTVSCRILTRPSPEFDFYLVPPPTVESKTKPFTLPLVLTGELSLS 240
 DB 181 MLYTPLRSNGSGDDVFTVSCRILTRPSPEFDFYLVPPPTVESKTKPFTLPLVLTGELSLS 240
 QY 241 RFPPLSIDEMVTSNPESIVQFQNGRVTLTGELTGTQLOACNICSIRGKVTGQVPSQHM 300
 DB 241 RFPVPIQMTVSPNEVISVQCGNGRCCTLDGELQTTQLOVSGICAFKGEVTAHLHNDHL 300
 QY 301 WNLLEITNLNGTQFDDVPAPLGPDPFAGEVFGVLSQRNR----GESNPNARAHDAVVA 356
 DB 301 NNVTITNLNGSPFDDVPAPLGPDPFAGEVFGVLSQRNR----GESNPNARAHDAVVA 356
 QY 357 TYSKYTPKGLVQIGTWTNTDVE-NQPTKFTPTIGLNEVANGHREQWTLPRYSGALTIN 415
 DB 361 TYTAQYTPKGLQIGTWTNTDVTNQNQPKFTPVGLNDT---EHFNQWVPRYAGALNIN 417
 QY 416 MNLAPAVAPLPGERLLFFRSYVPLKGGFGNPAIDCSVPQEWQVHFEQESAPSLGDVALV 475
 DB 418 TNLAPSVAVPFGERLLFFRSYVPLKGGYGNPAIDCLLPQGWQVHFEQEAAPSMSEVALV 477
 QY 476 RYVNPDTGRVLFEAKLHKGGLTIVSSSTSTGVPVVPANGYKFDPSWVNOFYSLAPMGGTNG 535
 DB 478 RYINPDGRALFEAKLHKGGLTIVSSSTSTGVPVVPANGYKFDPSWVNOFYSLAPMGGTNG 537
 QY 536 RRRVQ 540
 DB 538 RRRVQ 542
 RESULT 10
 Q915C7 PRELIMINARY; PRT; 548 AA.
 ID Q915C7
 AC Q915C7;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Capsid protein.
 OS Human calicivirus NLV/Erfurt/546/00/DB.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 OX NCBI_TaxID=173929;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NLV/Erfurt/546/00/DB;
 RX Kuenkel U., Hoehne M., Schreier E.;
 RA "Molecular epidemiology of outbreaks of gastroenteritis associated
 RT with Norwalk-like viruses in Germany";
 RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF427118; AAL18874.1; -;
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_Cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 SQ SEQUENCE 548 AA; 59874 MW; 9A1D8BEE58ADE7D CRC64;
 Query Match 81.8%; Score 2326; DB 12; Length 548;
 Best Local Similarity 79.2%; Pred. No. 8.9e-176;
 Matches 435; Conservative 51; Mismatches 53; Indels 10; Gaps 6;
 QY 1 MKMASNDATPSNDGAAGLVPESSNEAMALEPVPVVGASLAAPVTGQTNIIIDPWIRTNFVQAP 60
 DB 1 MKMASNDAASTDGAAGLVPESSNEAMALEPVPVVGASLAAPVTGQTNIIIDPWIRTNFVQAP 60
 QY 61 NGEFTVSPRNSPGEIILVNLELGPENLPYLAHARMYNGVAGMEVQVMLAGNAFTAGKII 120
 DB 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHARMYNGVAGMEVQVMLAGNAFTAGKII 120
 QY 121 FAAPVPFPVENLSPSQITMFPFHVIIIDVRLPEVLLPMPDVRSTLPHNQDEPKRLVA 180
 DB 121 FAAPVPFPVENLSPSQITMFPFHVIIIDVRLPEVLLPMPDVRSTLPHNQDEPKRLVA 180
 QY 181 MLYTPLRSNGSGDDVFTVSCRILTRPSPEFDFYLVPPPTVESKTKPFTLPLVLTGELSLS 240
 DB 181 MLYTPLRSNGSGDDVFTVSCRILTRPTDFDFYLVPPPTVESKSKPFTLPLVLTGELSLS 240
 QY 241 RFPPLSIDEMVTSNPESIVQFQNGRVTLTGELTGTQLOACNICSIRGKVTGQVPSQHM- 299
 DB 241 RFPPLSIDVLYNPNESAIVQCGNGRCCTLDGELQTTQLOVSGICAFKGEVTAHLHNDHL 300
 QY 300 -MNLLEITNLNGTQFDDVPAPLGPDPFAGEVFGVLSQRNR----GESN-PANRAHDA 353
 DB 301 THWNTVITNLNGTQFDDVPAPLGPDPFAGEVFGVLSQRNR----GESN-PANRAHDA 360
 QY 354 VVATYSKYTPKGLVQIGTWTNTDVE-NQPTKFTPTIGLNEV-ANGHREQWTLPRYSGA 411
 DB 361 VIATYSKFTPKLGNIQFSTWETQDVSSGQPTKFTTPVGLASVDANSH-FDQWTLPSYGA 419
 QY 412 LTLNMLAPAVAPLPGERLLFFRSYVPLKGGFGNPAIDCSVPQEWQVHFEQESAPSLGD 471
 DB 420 LTLNMLAPSVAVPFGERLLFFRSYVPLKGGYGNPAIDCLMPQEWQVHFEQESAPSLSD 479
 QY 472 VALVRYNPDTGRVLFEAKLHKGGLTIVSSSTSTGVPVVPANGYKFDPSWVNOFYSLAPMG 531
 DB 480 VALVRYNPDETGRVLFEAKLHKGGLTIVARNAGVPAVPTNGYPRFDSWVNOFYTLAPMG 539
 QY 532 TNGRRVQ 540
 DB 540 NGSGRRRMQ 548
 RESULT 11
 Q91115 PRELIMINARY; PRT; 535 AA.
 ID Q91115
 AC Q91115;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Capsid protein.
OS Human calicivirus NLV/Schwerin 003/00/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
NCBI_TaxID=168852;
RN [1]_TaxID=168852;
RP SEQUENCE FROM N.A.
RA Kuenkel U., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated
with Norwalk-like viruses in Germany."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF397905; AAK92128.1;
DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF00915; Calici_coat_1.
SQ SEQUENCE 535 AA; 58520 MW; 2E97BBD3F3D5FCC1 CRC64;

Query Match 79.9%; Score 2273.5; DB 12; Length 535;
Best Local Similarity 76.5%; Pred. No. 1.2e-171;
Matches 414; Conservative 65; Mismatches 55; Indels 7; Gaps 3;

QY 1 MKMASNDATPSNDGAGLVPESSNNEAMALEPVVGASLAAPVTGQTNIIIDPWIRTNFVQAP 60
DB 1 MKMASNDATPSNDGAGLVPESSNNEAMALEPVVGASLAAPVTGQTNIIIDPWIRTNFVQAP 60
QY 61 NGEFTVSPRNSPGEILVNLELGPENLPYLAHLARMYNGYAGMEVQVMLAGNAFTAGKII 120
DB 61 NGEFTVSPRNSPGEILVNLELGPENLPYLAHLARMYNGYAGMEVQVMLAGNAFTAGKII 120
QY 121 FAAPVPYPPVENLSPQITMFPVHVIDVRLTPVLLPMDVRLSTLPHFNQKDEPKRLVA 180
DB 121 FAAPVPYPPVENLSPQITMFPVHVIDVRLTPVLLPMDVRLSTLPHFNQKDEPKRLVA 180
QY 181 MLYTPLRNSGDDVFTVSCRILTRPSPEFDTYLVPPTVESKTKPFTLPVLTIGELNS 240
DB 181 MLYTPLRNSGDDVFTVSCRILTRPSPEFDTYLVPPTVESKTKPFTLPVLTIGELNS 240
QY 241 RPPLSIDEMVTSNPSISVVPQNGRVTLDGELLGTTQLOACNICISIRGKVTQVPSQOHM 300
DB 241 RPPLSIDEMVTSNPSISVVPQNGRVTLDGELLGTTQLOACNICISIRGKVTQVPSQOHM 300
QY 301 WNLEITNLNGTQFDPTDDVPAPLGVDPFAGEVGVLSQRNRESNPANRAHDVAATYSD 360
DB 301 WNLEITNLNGTQFDPTDDVPAPLGVDPFAGEVGVLSQRNRESNPANRAHDVAATYSD 360
QY 361 KYTPKGLGVQIGTWNNDVE--NPTKFTPIGLNEVANGHRFEQWTLPRYSGALTINMNL 419
DB 361 KYTPKGLGVQIGTWNNDVE--NPTKFTPIGLNEVANGHRFEQWTLPRYSGALTINMNL 419
QY 419 PPVAPTFGEQILFFRSHIPLAGGVADPVIDCLLQEMIQHLYQESAPSQSDVALIRTN 474
DB 419 PPVAPTFGEQILFFRSHIPLAGGVADPVIDCLLQEMIQHLYQESAPSQSDVALIRTN 474
QY 480 PDTGRVLFPAKLHKGGLTVSSTGTPVVVPANGYFKPDSWNQYSLAPMGTGNRRRV 539
DB 480 PDTGRVLFPAKLHKGGLTVSSTGTPVVVPANGYFKPDSWNQYSLAPMGTGNRRRV 539
QY 475 PDTGRVLFPAKLHRSQITVANTGSRPIVVPANGYFRFDSWNQYSLAPMGTGNRRRV 534
DB 475 PDTGRVLFPAKLHRSQITVANTGSRPIVVPANGYFRFDSWNQYSLAPMGTGNRRRV 534
QY 540 Q 540
DB 535 Q 535

RESULT 12
Q68104 PRELIMINARY; PRT; 535 AA.
AC Q68104;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Hawaii calicivirus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.

OX NCBI_TaxID=33750;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;
RX MEDLINE=94358485; PubMed=8077710;
RA Lew J.F., Kapikian A.Z., Valdeusio J., Green K.Y.;
RT "Molecular characterization of Hawaii virus and other Norwalk-like
viruses: evidence for genetic polymorphism among human
caliciviruses."
RL J. Infect. Dis. 170:535-542(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;
RA Lew J.F.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;
RA Green K.Y., Sosnovtseva S.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;
RA Pletneva M.A., Sosnovtseva S., Green K.Y.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: U07611; AAB97768.2;
DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF00915; Calici_coat_1.
SQ SEQUENCE 535 AA; 58741 MW; DOCA252A210E4599 CRC64;

Query Match 79.8%; Score 2270.5; DB 12; Length 535;
Best Local Similarity 76.3%; Pred. No. 2.2e-171;
Matches 413; Conservative 63; Mismatches 58; Indels 7; Gaps 3;

QY 1 MKMASNDATPSNDGAGLVPESSNNEAMALEPVVGASLAAPVTGQTNIIIDPWIRTNFVQAP 60
DB 1 MKMASNDATPSNDGAGLVPESSNNEAMALEPVVGASLAAPVTGQTNIIIDPWIRTNFVQAP 60
QY 61 NGEFTVSPRNSPGEILVNLELGPENLPYLAHLARMYNGYAGMEVQVMLAGNAFTAGKII 120
DB 61 NGEFTVSPRNSPGEILVNLELGPENLPYLAHLARMYNGYAGMEVQVMLAGNAFTAGKII 120
QY 121 FAAPVPYPPVENLSPQITMFPVHVIDVRLTPVLLPMDVRLSTLPHFNQKDEPKRLVA 180
DB 121 FAAPVPYPPVENLSPQITMFPVHVIDVRLTPVLLPMDVRLSTLPHFNQKDEPKRLVA 180
QY 181 MLYTPLRNSGDDVFTVSCRILTRPSPEFDTYLVPPTVESKTKPFTLPVLTIGELNS 240
DB 181 MLYTPLRNSGDDVFTVSCRILTRPSPEFDTYLVPPTVESKTKPFTLPVLTIGELNS 240
QY 241 RPPLSIDEMVTSNPSISVVPQNGRVTLDGELLGTTQLOACNICISIRGKVTQVPSQOHM 300
DB 241 RPPLSIDEMVTSNPSISVVPQNGRVTLDGELLGTTQLOACNICISIRGKVTQVPSQOHM 300
QY 301 WNLEITNLNGTQFDPTDDVPAPLGVDPFAGEVGVLSQRNRESNPANRAHDVAATYSD 360
DB 301 WNLEITNLNGTQFDPTDDVPAPLGVDPFAGEVGVLSQRNRESNPANRAHDVAATYSD 360
QY 361 KYTPKGLGVQIGTWNNDVE--NPTKFTPIGLNEVANGHRFEQWTLPRYSGALTINMNL 419
DB 361 KYTPKGLGVQIGTWNNDVE--NPTKFTPIGLNEVANGHRFEQWTLPRYSGALTINMNL 419
QY 419 PPVAPTFGEQILFFRSHIPLAGGVADPVIDCLLQEMIQHLYQESAPSQSDVALIRTN 474
DB 419 PPVAPTFGEQILFFRSHIPLAGGVADPVIDCLLQEMIQHLYQESAPSQSDVALIRTN 474
QY 480 PDTGRVLFPAKLHKGGLTVSSTGTPVVVPANGYFKPDSWNQYSLAPMGTGNRRRV 539
DB 480 PDTGRVLFPAKLHKGGLTVSSTGTPVVVPANGYFKPDSWNQYSLAPMGTGNRRRV 539
QY 475 PDTGRVLFPAKLHRSQITVANTGSRPIVVPANGYFRFDSWNQYSLAPMGTGNRRRV 534
DB 475 PDTGRVLFPAKLHRSQITVANTGSRPIVVPANGYFRFDSWNQYSLAPMGTGNRRRV 534
QY 540 Q 540

Db	535 Q 535	Db	535 Q 535
RESULT 14			
Q915C6	PRELIMINARY; PRT; 535 AA.	Q915C6	PRELIMINARY; PRT; 535 AA.
AC	Q915C6	AC	Q915C6
DT	01-DEC-2001 (Tremblrel. 19, Created)	DT	01-DEC-2001 (Tremblrel. 19, Created)
DT	01-MAY-2000 (Tremblrel. 13, Last sequence update)	DT	01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)	DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE	Capsid protein.	DE	Capsid protein.
OS	Human calicivirus NLV/Pirna/110/00/DE.	OS	Human calicivirus NLV/Pirna/110/00/DE.
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;	OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC	Norovirus.	OC	Norovirus.
OX	NCBI_TaxID=173930;	OX	NCBI_TaxID=173930;
RN	[1]	RN	[1]
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RC	STRAIN=NLV/Pirna/110/00/DE;	RC	STRAIN=NLV/Pirna/110/00/DE;
RA	Kuenkel U., Hoehne M., Schreier E.;	RA	Kuenkel U., Hoehne M., Schreier E.;
RT	Molecular epidemiology of outbreaks of gastroenteritis associated	RT	Molecular epidemiology of outbreaks of gastroenteritis associated
RL	with Norwalk-like viruses in Germany."	RL	with Norwalk-like viruses in Germany."
DR	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.	DR	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF427119; AAL18875.1; -	DR	EMBL; AF427119; AAL18875.1; -
DR	InterPro; IPR004005; Calici_coat.	DR	InterPro; IPR004005; Calici_coat.
DR	InterPro; IPR008975; Viral_cap_coat.	DR	InterPro; IPR008975; Viral_cap_coat.
DR	Pfam; PF00915; Calici_coat; 1.	DR	Pfam; PF00915; Calici_coat; 1.
DR	SEQUENCE 535 AA; 58505 MW; D839C359F86DC1F9 CRC64;	DR	SEQUENCE 535 AA; 58505 MW; D839C359F86DC1F9 CRC64;
Query Match 79.8%; Score 2269.5; DB 12; Length 535;			
Best Local Similarity 76.3%; Pred. No. 2.6e-171;			
Matches 413; Conservative 65; Mismatches 56; Indels 7; Gaps 3;			
Qy	1 MKMASNDATPSNDGAAGLVPSNEAMALEPVVVGASLAAPVTGQTNIIDPWIRTNFVQAP 60	Qy	1 MKMASNDATPSNDGAAGLVPSNEAMALEPVVVGASLAAPVTGQTNIIDPWIRTNFVQAP 60
Db	1 MKMASNDAAASNDGAAGLVPEANNEITMALEPVAGASIAAPLTGQNNVIDPWIRNFVQAP 60	Db	1 MKMASNDAAASNDGAAGLVPEANNEITMALEPVAGASIAAPLTGQNNVIDPWIRNFVQAP 60
Qy	61 NGFTVSPRSPGILVNLELGPENLPYLAHLARMYNGYAGGVEQVLLAGNAFTAGKII 120	Qy	61 NGFTVSPRSPGILVNLELGPENLPYLAHLARMYNGYAGGVEQVLLAGNAFTAGKII 120
Db	61 NGFTVSPRSPGILVNLELGPENLPYLAHLARMYNGYAGGVEQVLLAGNAFTAGKLV 120	Db	61 NGFTVSPRSPGILVNLELGPENLPYLAHLARMYNGYAGGVEQVLLAGNAFTAGKLV 120
Qy	121 FAAPVPPFVPEVNLSPSQITMPPHVIIDVRLPEVLLPMPDVRSTLHFNQDKPQMLVA 180	Qy	121 FAAPVPPFVPEVNLSPSQITMPPHVIIDVRLPEVLLPMPDVRSTLHFNQDKPQMLVA 180
Db	121 FAAPVPPFVPEVNLSPSQITMPPHVIIDVRLPEVLLPMPDVRSTLHFNQDKPQMLVA 180	Db	121 FAAPVPPFVPEVNLSPSQITMPPHVIIDVRLPEVLLPMPDVRSTLHFNQDKPQMLVA 180
Qy	181 MLYTPLRNSGSDVFTVSCRLTRPSDFDFTVLPPTVESKTKPFTLPIITIGELTNS 240	Qy	181 MLYTPLRNSGSDVFTVSCRLTRPSDFDFTVLPPTVESKTKPFTLPIITIGELTNS 240
Db	181 MLYTPLRNSGSDVFTVSCRLTRPSDFDFTVLPPTVESKTKPFTLPIITIGELTNS 240	Db	181 MLYTPLRNSGSDVFTVSCRLTRPSDFDFTVLPPTVESKTKPFTLPIITIGELTNS 240
Qy	241 RPLSIDEMVTSNPESIIVQPNQGRVTLDGELLGTTLQACNICSIKQKVTGVPSQHM 300	Qy	241 RPLSIDEMVTSNPESIIVQPNQGRVTLDGELLGTTLQACNICSIKQKVTGVPSQHM 300
Db	241 RFPAPFIDELTSPNESLVVQPNQGRVTLDGELLGTTLQACNICSIKQKVTGVPSQHM 300	Db	241 RFPAPFIDELTSPNESLVVQPNQGRVTLDGELLGTTLQACNICSIKQKVTGVPSQHM 300
Qy	301 WNELETNLNGTQFDPTDDVPAPLGVDPFAGVFGVLSQNRNGESNPANRAHDVAIVATYSD 360	Qy	301 WNELETNLNGTQFDPTDDVPAPLGVDPFAGVFGVLSQNRNGESNPANRAHDVAIVATYSD 360
Db	301 WNMQVNTNGTQFDPTDDVPAPLGVDPFAGVFGVLSQNRNGESNPANRAHDVAIVATYSD 360	Db	301 WNMQVNTNGTQFDPTDDVPAPLGVDPFAGVFGVLSQNRNGESNPANRAHDVAIVATYSD 360
Qy	361 KYTPKGLVQIGTWNNDVE-NQPTKFTPIGLNEVANGHREQWTLPRYSGALTINMGLA 419	Qy	361 KYTPKGLVQIGTWNNDVE-NQPTKFTPIGLNEVANGHREQWTLPRYSGALTINMGLA 419
Db	361 KYTPKGLVQIGTWNNDVE-NQPTKFTPIGLNEVANGHREQWTLPRYSGALTINMGLA 419	Db	361 KYTPKGLVQIGTWNNDVE-NQPTKFTPIGLNEVANGHREQWTLPRYSGALTINMGLA 419
Qy	420 PAVAPLFGGERLLFRSVVPLKGGFNGNPAIDCSVPQEWQHFQESAPSLGDVALVRVN 479	Qy	420 PAVAPLFGGERLLFRSVVPLKGGFNGNPAIDCSVPQEWQHFQESAPSLGDVALVRVN 479
Db	415 PPVAPTFPGEQILFRSHIPLKGGVADPVIDCLLPQEWIQLHYQESAPSQSDVALIRFTN 474	Db	415 PPVAPTFPGEQILFRSHIPLKGGVADPVIDCLLPQEWIQLHYQESAPSQSDVALIRFTN 474
Qy	480 PDTGRVLPFAKLHKGGLTVSSTSTGVPVVPANGYFKFDSWVNOFYSLAPMGTGNGRRRV 539	Qy	480 PDTGRVLPFAKLHKGGLTVSSTSTGVPVVPANGYFKFDSWVNOFYSLAPMGTGNGRRRV 539
Db	475 PDTGRVLPFAKLHRSYITVANTGSRPIVVPANGYFRFDSWVNOFYSLAPMGTGNGRRRV 534	Db	475 PDTGRVLPFAKLHRSYITVANTGSRPIVVPANGYFRFDSWVNOFYSLAPMGTGNGRRRV 534
Qy	540 Q 540	Qy	540 Q 540
Db	535 Q 535	Db	535 Q 535

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:36:02 ; Search time 10.323 Seconds
(without alignments)
4985.230 Million cell updates/sec

Title: US-09-926-799-8
Perfect score: 2843
Sequence: 1 MKMASNDAPSDGAGLVLP.....VNQFVSLAPMTGNGRRRVQ 535

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2298.5	80.8	542	S60616	capsid protein - h
2	1923	67.6	539	S40111	capsid protein - h
3	1196.5	42.1	530	B37471	capsid protein - N
4	1174	41.3	546	B37491	major capsid prote
5	328	11.5	576	A53982	capsid protein - E
6	327.5	11.5	2344	S55399	genome polyprotein
7	322.5	11.3	2344	RRWRH	genome polyprotein
8	314.5	11.1	2344	S64740	genome polyprotein
9	289	10.2	702	A48562	coat protein - San
10	284	10.0	668	VCWVFC	coat protein - fel
11	277.5	9.8	668	VCW354	capsid protein - f
12	277.5	9.8	671	VCWVF9	coat protein - fel
13	273.5	9.6	703	A48562	coat protein - San
14	272.5	9.6	668	VCWVFF	coat protein - fel
15	266	9.4	668	VCW356	capsid protein - f
16	130	4.6	2206	GNNY27	genome polyprotein
17	128	4.5	2194	GNNYE7	genome polyprotein
18	127	4.5	733	VCW354	capsid protein - f
19	126	4.4	733	VCW356	capsid protein - f
20	125.5	4.4	2333	GNNY2F	genome polyprotein
21	122	4.3	378	E96787	protein T402.5 [i
22	121.5	4.3	2332	GNNYF	genome polyprotein
23	120.5	4.2	2205	GNNY2W	genome polyprotein
24	120	4.2	1011	GNNYC1	genome polyprotein
25	119.5	4.2	2336	S37077	genome polyprotein
26	115.5	4.1	757	JS0198	genome polyprotein
27	115	4.0	2207	S09553	genome polyprotein
28	115	4.0	2332	GNNY4F	genome polyprotein
29	114	4.0	2206	S03822	genome polyprotein

30	114	4.0	2207	1	GNNYSP	genome polyprotein
31	112	3.9	2206	1	GNNY4P	genome polyprotein
32	110.5	3.9	2468	2	AB3412	hypothetical prote
33	110	3.9	716	2	S45262	NF-AT component -
34	109	3.8	833	2	S45041	genome polyprotein
35	109	3.8	913	2	JG0168	gob-5 protein - mo
36	109	3.8	1072	2	S76888	hypothetical prote
37	109	3.8	2201	1	GNNYA9	genome polyprotein
38	108.5	3.8	3262	2	AH2137	hypothetical prote
39	107.5	3.8	929	2	A44048	genome polyprotein
40	107	3.8	703	2	H86349	protein F8K7.9 [im
41	106.5	3.7	3085	2	T00327	polyprotein - infe
42	102.5	3.6	800	2	H84740	hypothetical prote
43	102.5	3.6	1698	2	T13800	coracle gene prote
44	102	3.6	2629	2	I46569	apolipoprotein B -
45	101.5	3.6	852	1	GNNYHA	genome polyprotein

ALIGNMENTS

RESULT 1

S60616
capsid protein - human calicivirus (strain Melksham)
C:Species: human calicivirus

A:Variety: strain Melksham
C:Date: 23-May-1997 #sequence_revision 23-May-1997 #text_change 28-Jul-2000

C:Accession: S60616
R:Green, S.M.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.

Virus Res. 37, 271-283, 1995

A:Title: Capsid diversity in small round-structured viruses: molecular characterization
A:Reference number: S60615; MUID:96136658; PMID:9533462
A:Accession: S60616
A:Molecule-type: Genomic RNA

A:Residues: 1-542 <GRE>

A:Cross-references: ENBL:X81879; NID:G976077; PIDN:CAA57462.1; PID:G976079
A:Experimental source: strain Melksham
A>Note: it is uncertain whether Met-1 or Met-3 is the Initiator

C:Superfamily: human calicivirus capsid protein

C:Keywords: capsid protein; coat protein

Query Match 80.8%; Score 2298.5; DB 2; Length 542;

Best Local Similarity 76.6%; Pred. No. 7.3e-162;

Matches 415; Conservative 65; Mismatches 55; Indels 7; Gaps 1;

QY	1	MMKASNDAPSDGAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP	60
DB	1	MMKASNDAPSDTGAAGLVPESSNEVMALEPVAGAAALAPVTGQTNIIDPWIRLNFVQAP	60
QY	61	NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMYNGYAGGVGVQVLLAGNAFTAGKLV	120
DB	61	NGEFTVSPRNAPGEVLLNLELGPENLPYLAHLARMYNGYAGGVGVQVLLAGNAFTAGKLV	120
QY	121	FAAVPPHPFLENISPGQITMFPVHVIDVRLPLEPVLPLPDVNRNFFHYNQNEPRMLVA	180
DB	121	FAAVPPHPFVENLSPGQITMFPVHVIDVRLPLEPVLPLPDVNRNFFHYNQNEPRMLVA	180
QY	181	MLYTLRNSGSGDDVFTVSCRVLTRPSDFDFNVLVPTLESKTKPFTLPITIGELTNS	240
DB	181	MLYTLRNSGSGDDVFTVSCRVLTRPSDFDFNVLVPTLESKTKPFTLPITIGELTNS	240
QY	241	RFPVPIDELTSPNBSLVVQPNQRCALDGLQGTQTLPLTAICSFGRINOKVSGENHV	300
DB	241	RFPVPIDQMTSPNEVISVQCNQGRCTLDGLQGTQTLQVSGICAFKGEVTAHLHDNDHL	300
QY	301	WNQVNTNNGPFDPTGDPAPLPTDFSGKLVGVLSORDHDNAC-----RSHDAVTA	353
DB	301	NNVTITNLNSPFDPSDIPAPLGVDFEQGRVFGVISQRDKQNAAGHSEPNRGRHDVVP	360
QY	354	TNSAKFTKLGAIQIGTWEDDVHINQPTKFTPVGLFENEGFNQWTLPNYSGALTINMGL	413
DB	361	TYTAQYTKLGAIQIGTWQTDLTNTVNPQVKFTPVGLNDEHFNQWVPRVYAGALNLTNL	420

414 APPVAPTPGEQILFFRSHIPLKGGVADPVIDCLLPQEWIQLHLYQESAPSDVALIRFT 473
 421 ABSVAPVFGERLLFFRSHLPLKGGVGNPAIDCLLPQEWVQHFYQEAAPSSEVALVRYI 480
 474 NPDTGRVLFEAKLHRSYITVANTGSRPLVVPANGYFRFDTWVNOFYSLAPMGVGNRR 533
 481 NPDTGRALFEAKLHRSYITVANTGSRPLVVPANGYFRFDTWVNOFYSLAPMGVGNRR 540
 534 VQ 535
 541 VQ 542

RESULT 2
 S40111
 capsid protein - human calicivirus (strain Bristol isolate B493)
 C:Species: human calicivirus
 A:Variety: Strain Bristol isolate B493
 C:Date: 25-Dec-1994 #sequence_revision 27-Feb-1997 #text_change 28-Jul-2000
 C:Accession: S40111
 R:Green, S.M.; Dingle, K.E.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
 submitted to the EMBL Data Library, December 1993
 A:Description: Human enteric caliciviridae: a new prevalent SRSV group defined by RNA-de
 A:Reference number: S40111
 A:Accession: S40111
 A:Molecule type: genomic RNA
 A:Residues: 1-539 <GRE>
 A:CROSS-references: EMBL:X76716; NID:G436410; PIDN:CAA54134.1; PID:G436411
 A:Experimental source: human enteric calicivirus strain Bristol isolate B493
 C:Superfamily: human calicivirus capsid protein
 C:Keywords: capsid protein; coat protein

Query Match 67.6%; Score 1923; DB 2; Length 539;
 Best Local Similarity 65.7%; Pred. No. 3.8e-134;
 Matches 355; Conservative 76; Mismatches 97; Indels 10; Gaps 5;
 1 MKMASNDAPNSDGAAGLVPEANETMALEPVAGASIAAPLTGONNIIDPWIRLNFOVAP 60
 1 MKMASNDANPSDGAANLVEVNEVNALEPVVGAALAAEVAGQVQVVDPIWIRNFOVAP 60
 61 NGFTVSPRNSPGBVLNLLELGPENLPYLAHLSRMVYAGGVQVOLLAGNAFTAGKLV 120
 61 GGETVSPRNPAGEILMSAPLGPDLNPLSHLSRMVYAGGVQVOLLAGNAFTAGKVI 120
 121 FAAPVPPHPLNISPQGITMFPFHVIVDRTLEPVLPLPDVRNFFHYNQNEPRMLVA 180
 121 FAAPVPPHPLNISPQGITMFPFHVIVDRTLEPVLPLPDVRNFFHYNQNEPRMLVA 180
 181 MLYTPLRNSGDDVFTVSCRVLTSPDFFDENVLPVPTLESKTKPTLPILTIGELTNS 240
 181 MLYTPLRANAGDDVFTVSCRVLTSPDFFDENVLPVPTLESKTKPTLPILTIGELTNS 240
 241 RFPVPIDELTYSPNESLVVQPNQRCALDGELOQTTLQPLTAICSPFGRINQKVSQENHV 300
 241 RFPVPIDELTYSPNESLVVQPNQRCALDGELOQTTLQPLTAICSPFGRINQKVSQENHV 300
 301 WNMQVTNITNGTPTDGVDPAPLGTDPFSGKLVGLVSRQDH-DNACSHDAVIATNSAKF 359
 299 YTNLASQWNSNDYPTIEPAPLGTDPFSGKLVGLVSRQDH-DNACSHDAVIATNSAKF 358
 360 TPKGAIQIGTWEEEDVHINOPTKFTPVGLF-----ENEGFNQWTLNYSGLATLNMGL 413
 359 TPKGVSQVTTDTNNDFOAGQNTKFTPVGLVQDGDHQNKP-QQWLLPNYSGRTHNVHL 417
 414 APPVAPTPGEQILFFRSHIPLKGGVADPVIDCLLPQEWIQLHLYQESAPSDVALIRFT 473
 418 APAVAPTPGEQILFFRSHIPLKGGVADPVIDCLLPQEWIQLHLYQESAPSDVALIRFT 477
 474 NPDTGRVLFEAKLHRSYITVANTGSRPLVVPANGYFRFDTWVNOFYSLAPMGVGNRR 533
 478 NPDTGRVLFECKLHRSYITVAHTGPDYDLVLPNGYFRFDTWVNOFYSLAPMGVGNRR 537

RESULT 3
 B37471
 capsid protein - Norwalk virus
 C:Species: Norwalk virus
 C:Date: 24-Feb-1994 #sequence_revision 25-Apr-1997 #text_change 28-Jul-2000
 C:Accession: B37471
 R:Jiang, X.; Wang, M.; Wang, K.; Estes, M.K.
 Virology 195, 51-61, 1993
 A:Title: Sequence and genomic organization of Norwalk virus.
 A:Reference number: A37471; MUID:93303939; PMID:8391187
 A:Accession: B37471
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: genomic RNA
 A:Residues: 1-530 <JIA>
 A:CROSS-references: GB:M87661; NID:G1061311; PIDN:AAB50466.1; PID:G1061313
 A:Note: sequence extracted from NCBI backbone (NCBIP:134157)
 C:Superfamily: human calicivirus capsid protein

Query Match 42.1%; Score 1196.5; DB 2; Length 530;
 Best Local Similarity 46.4%; Pred. No. 1.6e-80;
 Matches 254; Conservative 81; Mismatches 172; Indels 41; Gaps 10;
 1 MKMASNDAPNSDGAAG---LVPEAN-NEETMALEPVAGASIAAPLTGONNIIDPWIRLN 56
 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWINNF 60
 57 VOAPNGEFTVSPRNSPGBVLNLLELGPENLPYLAHLSRMVYAGGVQVOLLAGNAFTA 116
 61 VOAPQGEFTISPNTPGDVLFDSLGLPHLNPFLHLSQMYNGVGMVRVIMLAGNAFTA 120
 117 GKLVAAPVPPHPLNISPQGITMFPFHVIVDRTLEPVLPLPDVRNFFHYNQNEPRM 176
 121 GKIIVSCIPPGFGSHNLTAQATLFPFHVIVDRTLEPVLPLPDVRNFFHYNQNEPRM 180
 177 RLVAHMLTYPLRNSGDDVFTVSCRVLTSPDFFDENVLPVPTLESKTKPTLPILTIGE 236
 181 RLVCMLTYPLRNSGDDVFTVSCRVLTSPDFFDENVLPVPTLESKTKPTLPILTIGE 240
 237 LTNSRFPVPIDELTYSPNESLVVQPNQRCALDGELOQTTLQPLTAICSPFGRINQKVS 296
 241 LNSRFPVPIDELTYSPNESLVVQPNQRCALDGELOQTTLQPLTAICSPFGRINQKVS 298
 297 ENHVMQVTNITNGTPTDGVDPAPLGTDPFSG-----KLVGLVSRQDHDAVIATNSAKF 349
 299 -----INLTDELDTGTPFHP-EGPAPITGPDGCDHINNTQGHSSQTYD----- 344
 350 AVIATNSAKFTPKLGAIO---IGTWEEEDVHINOPTKFTPVGLFENEGFNQWTLNYSGA 406
 345 --VDITPTDTPFVHLGSIQANGISGN----YVGLVSRQDHDAVIATNSAKF 398
 407 LTNMGLAPPVAPTPGEQILFFRSHIPLKGGVADPVIDCLLPQEWIQLHLYQESAPSD 466
 399 ITEATHLAPSVVPPGGEVLFVFMSPKMPGAYNLP---CLLPQEWIQLHLYQESAPSD 455
 467 VALIRFTPDTCRVLFEAKLHRSYITVANTG--SRPVTVPANGYFRFDTWVNOFYSLAP 524
 456 AALLHYVDPTGRNLGFKAYPDGFLCTVCPNGASSGPQQLPGLINGVVFVSVSRFYQLKP 515
 525 MGTGNGRR 532
 516 VGTASSAR 523

RESULT 4
 B37491
 major capsid protein [similarity] - Southampton virus
 N:Alternate names: orf2 protein
 C:Species: Southampton virus
 C:Date: 22-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 18-Aug-2000
 C:Accession: B37491
 R:Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
 Science 259, 516-519, 1993
 A:Title: Sequence and genome organization of a human small round-structured (Norwalk-like

A:Reference number: A37491; MUID:93142023; PMID:8380940

A:Accession: B37491

A>Status: not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-546 <L&M>

A:Cross-references: GB:L07418; NID:q1236787; PIDN:AAA92984.1; PID:g295114

A>Note: sequence extracted from NCBI backbone (NCBI:123456)

A:Title: small round-structured virus, SRSV, Norwalk virus, Norwalk-like virus, serotype

C:Superfamily: human calicivirus capsid protein

C:Keywords: glycoprotein

F:303,340,441/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.3%; Score 1174; DB 2; Length 546;

Best Local Similarity 46.2%; Pred. No. 7.6e-79;

Matches 261; Conservative 82; Mismatches 172; Indels 50; Gaps 16;

QY 1 MMASNDAAPSNDGAAG--LVPEANN--ETMALEPVAGASIAAPLTGQNNIIDPWIRLNF 56

DB 1 MMASKDAPOSADGASGAGQLVPEVNTADPLMEPVAGTTAVATAGQNMIDPWVNF 60

QY 57 VOAPNGEFTVSPRNSPGEVLLNLELGPENLPYLAHLRMYNGYAGGVEQVLLAGNAFTA 116

DB 61 VQSPQGEFTISPNTPGDIPLDLGLPHLPPLSHLSQMYNGWGMVRILLAGNAFSA 120

QY 117 GKLVFAAVPPHPLENIISPGQITMPFHVIIIVRTLEPVLPLPVDVRRNFFHYNQNEPRM 176

DB 121 GKIIIVCCVPPGPTSSSLTAQATLPHVIAVRTLEPIEMPLEDVRNVLVHTN-DNQPTM 179

QY 177 RLAVMLYPLRS-NGSGD-DVFTVSCRVLTRSPDPDFNLYVPPTLESKTKPFTLPILT 234

DB 180 RLVCMLYPLRTGGSGMSDSFVAVGRVLTAPESSDFLFLVPPTIEOKTRAFTVPIPL 239

QY 235 GELTNSRFPVPTIDELYTSNPSLVVQPNQRCALDGLQGTITQLLPTAICSPGRINQV 294

DB 240 QTLNSRFPVPTIDELYTSNPSLVVQPNQRCALDGLQGTITQLLPTAICSPGRINQV 299

QY 295 SCENHVMNQVNTINGTPDPDTPGVAPLPTDPSGKLVLSQRDHNAKSHDAV--- 351

DB 300 R-----TLNTEVDGKPF-MAFDSAPVGFDPF-GKCDWHMRISKTNNCTSGSDPMRSV 351

QY 352 -IATNSAKFTPKLGAIO-----IGTWEEDDVHINQPTKFTPVGLFENEGFNQW 398

DB 352 SVQTNVQGVFVPHLSIQFDEVNHTPTGYIGTIE-----WISQPS--TPPG-----TDINLW 401

QY 399 TLPNYSGALTLMNGLAPPVAPTFPGEQILFFRSHIPLKGVVADP-VIDCLLPQEWIOHLY 457

DB 402 ELPDYGSSLSQAANLAPPVFPFGCEALVYFVSAPFGPNRSAPNDVPCLLPQEVITHFV 461

QY 458 QESAPSDVALIRFTNPDTPGRVLPEAKLHRSYITVA--NTGSRPIVVPANGYFRFTW 515

DB 462 SEQAPTMGDAALLHVVDPTNRLGEFKLYPGGYLTCVPNGVGAGPQQLPLNGVFLFVSW 521

QY 516 VNOFYSLAPMGTCN-----GRRV 534

DB 522 VSRFYQLKPVGTASTARGRLGVRI 546

RESULT 5

AS3982

Capsid protein - European brown hare syndrome virus

C:Species: European brown hare syndrome virus

C:Title: European brown hare syndrome virus

C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 28-Jul-2000

C:Accession: A53982

R:Wirblich, C.; Meyers, G.; Ohlinger, V.F.; Capucci, L.; Ekenas, U.; Haas, B.; Thiel, H.

J. Virol. 68, 5164-5173, 1994

A:Title: European brown hare syndrome virus: relationship to rabbit hemorrhagic disease

A:Reference number: A53982; MUID:94309183; PMID:7518531

A:Accession: A53982

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-576 <WIR>

A:Cross-references: GB:U09199

C:Superfamily: human calicivirus capsid protein

Query Match

Best Local Similarity 11.5%; Score 328; DB 2; Length 576;

Matches 151; Conservative 77; Mismatches 160; Indels 186; Gaps 35;

QY 3 MASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAPNG 62

DB 31 VASTDVTADNVAAS-----VATAGIGGP-PQASPOESW-RVNFY--ND 72

QY 63 EFTVSPRNSPGEVLLNLELGPENLPYLAHLRMYNGYAGGVEQVLLAGNAFTAGKLVPA 122

DB 73 VFTSVTDAPGSILYVQHSQNNPFTQVLSQMYAGWAGMQFRIVAGSGIFGRLVCA 132

QY 123 AVPHFPLENTSPG-QITMFPHVIIIVRTLEPVLPLPVDVRRNFFHYNQNEPRM--RLV 179

DB 133 IIPPGI---QIQGLEVRQFPHVVIDARSLEPVTITMPDLRPEMYH--PTGDPGLVPTIV 187

QY 180 AMLYPLRSNGSGDDVFTVSCRVLTRSPDPDFNLYVPPTLESKT-----KP---ETLPL 232

DB 188 VSVNNL-INPFGGTTSAIQVTETRPSEDFEFLIRAPS--SKTVDSVNFSLWLTTPVL 244

QY 233 TIGELTNSRFPVPTIDELYTSNPSLVVQPNQRCALDGLQGTITQLLPTAICSPGRINQ 292

DB 245 T-GAGSDNRWCAPI-----VGLQVPFG----- 265

QY 293 KVGSENHVMNQVNTINGTPF-----DPTGDVPAPLPTDPSGKLVLSQRDH 341

DB 266 GFSTSNRHHNM-----NGSTYGSPPRDDIDHPSGNVSYPTGSATNTIETW-----Y 313

QY 342 DNACSHDAVIATNSAKFTPKLGAIOI-GTWEEDDVHINQPTKFTPVGLFENEGFNQ-WT 399

DB 314 ANAGTATTNPISNTAPDGFDPMGAIPFSGT-----TIPTGAW---VVGQGVWN 358

QY 400 LPN---YSGAL-TLNMGLA-----PVAPTFPGEQI-----LPFR 430

DB 359 ASNGTPYVGTVOAVELGPAFGANGAPSSIRPVTTT-TGAQLVAKSIYVATVATQNSSAGIIFL 417

QY 431 SHILPKGVADPVI--DCLLPQEWIOHLYQSSAPSDVALIRFTNPD-----GR--- 479

DB 418 S---KGMVSTPGVAATTYTFQ-----PS---AIV--TTPGTTPVAAPIGKNT 456

QY 480 VLPEAKLHRSYI-----TVANTGSRPIV 504

DB 457 IMFSAVVRTGDVNAVGPVNGTQVGVGSQPLSV 490

RESULT 6

S55399

genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)

A:Species: rabbit hemorrhagic disease virus

A:Variety: isolate BS89

C>Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 18-Jun-1999

C:Accession: S55399

R:Roessli, C.

submitted to the EMBL Data Library, May 1995

A:Reference number: S55399

A:Accession: S55399

A:Molecule type: genomic RNA

A:Residues: 1-2344 <ROS>

A:Cross-references: EMBL:X87607; NID:9854640; PIDN:CAA60910.1; PID:9854641

A:Experimental source: isolate BS89

C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase

C:Keywords: polyprotein

Query Match

Best Local Similarity 11.5%; Score 327.5; DB 2; Length 2344;

Matches 151; Conservative 74; Mismatches 226; Indels 125; Gaps 29;

QY 9 APSNDGAAGLVPEAN---NETMALEP-----VAGASIAAPLTQNNII 48

DB 1773 APOGE-AAGTATTASVPGTTTDDGLDGPVATTSSVVAENSSASITAGIGP-PQOVQ 1830

QY 49 DPWIRLNFVQAPNGEFTVSPRNSPGEVLLNLELGPENLPYLAHLRMYNGYAGGVEQVL 108

1831 ETW-RTNFY... 1887
109 LAGNAFTAGKLVFAA... 167
1888 VAGSGVFGRLVAA... 1944
168 YNOQNEPM--RLVAM... 224
1945 --PTGDPGLVPTLV... 1959
225 ----KP---FTLP... 270
2000 DISPAGLLTTPVLT... 2047
271 ELQTTQLLPTAIC... 330
2048 STYGWSSPRFAD... 2106
331 KLPGLVLSQRDHD... 388
2107 VPF-----NG... 2144
389 LFNENEGFNWLT... 448
2145 FATGAPGNLQPT... 2200
449 ----PQEWI... 495
2201 YTPQPDRIV... 2250
496 NTSGRPTW... 521
2251 GTGSOPLPV... 2286

RESULT 7
RRWRH
genome polyprotein - rabbit hemorrhagic disease virus
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C:Species: rabbit hemorrhagic disease virus
C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 29-May-1998
C:Accession: A41039
R:Mevers, G.; Wirblich, C.; Thiel, H.J.
Virology 184, 664-676, 1991
A:Title: Rabbit hemorrhagic disease virus--molecular cloning and nucleotide sequencing
A:Reference number: A41039; MUID:91361557; PMID:1840711
A:Accession: A41039
A:Molecule type: genomic RNA
A:Residues: 1-2344 <MEY>
A:Cross-references: GB:M67473
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: nucleotidyltransferase

Query Match 11.3%; Score 322.5; DB 1; Length 2344;
Best Local Similarity 26.5%; Pred. No. 4.5e-15;
Matches 154; Conservative 72; Mismatches 221; Indels 135; Gaps 31;

8 AAPSDGAAGLVPEAN--NETMALEP-----VAGASIAAPLTGQNNI 47
1772 AAPQGE-AAGTATTASVPGTTTDMGDPGVVATTVSITAENSASATAGIGP-PQQVDQ 1829
48 IDPWIRLNFVQAPNGEFTSPNSGCEVLNLELGPENLPYLHLRMYNGYAGGVEVQV 107
1830 QETW-RTNFY--NDVFTMSVADAPGSIITYVQHSFQNNPFTAVLSQMYAGWAGMQFRF 1886
108 LLAGNAFTAGKLVFAA... 166
1887 IVAGSGVFGRLVRAVIPP... 1943
167 HYNQNEPRM--RLVAM... 224
1944 H--PTGDPGLVPTLV... 1998

225 ----KP---FTLP... 273
1999 VDSISPAGLLTTPVLT... 2046
274 GTTQLLETA---ICS... 329
2047 GSTYQWSSPRFCD... 2105
330 GKLFGVLRSQRDHD... 387
2106 FVPF-----NG... 2143
388 GLFNENEGFNWLT... 447
2144 GFATGAPGNLQPT... 2195
448 LPOEWIQLHYQBSA... 492
2196 -----AQRIT... 2244
493 ---TVANTGSRPTW... 521
2245 ANGTOYGTSGQLPV... 2286

RESULT 8
S64740
genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)
N:Contains: VP60 protein
C:Species: rabbit hemorrhagic disease virus
A:Variety: isolate AST/89
C>Date: 12-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Jun-2000
C:Accession: S64740; S46944; S49018; S65012
R:Casais, R.; Martin-Alonso, J.; Boga, J.; Parra, F.
submitted to the EMBL Data Library, May 1995
A:Description: Genomic organization of rabbit hemorrhagic disease virus determined by di
A:Reference number: S64740
A:Accession: S64740
A:Molecule type: genomic RNA
A:Residues: 1-2344 <CAS>
A:Cross-references: EMBL:249271; NID:g1182032; PIDN:CAA89265.1; PID:g1150552
A:Experimental source: isolate AST/89
R:Boga, J.; Casais, R.; Martin, M.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra, F.
submitted to the EMBL Data Library, July 1993
A:Description: Molecular cloning, sequence and expression of the capsid protein gene fr
A:Reference number: S46944
A:Accession: S46944
A:Molecule type: genomic RNA
A:Residues: 1650-2344 <BOG>
A:Cross-references: EMBL:224757; NID:g515622; PIDN:CAA80881.1; PID:g515623
R:Parra, F.; Boga, J.A.; Martin, M.S.; Casais, R.
Virus Res. 27, 219-228, 1993
A:Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus sup
A:Reference number: S49018; MUID:93255896; PMID:8488721
A:Accession: S49018
A:Molecule type: genomic RNA
A:Residues: 1650-1796 <PAR>
A:Cross-references: EMBL:224757
A:Experimental source: isolate AST/89
A:Accession: S65012
A:Molecule type: protein
A:Residues: 1767-1779; 1875-1877; 'X', 1879-1881; 1936-1938; 'X', 1940-1941 <PAW>
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: blocked amino end; polyprotein

Query Match 11.1%; Score 314.5; DB 2; Length 2344;
Best Local Similarity 25.9%; Pred. No. 1.8e-14;
Matches 149; Conservative 75; Mismatches 227; Indels 125; Gaps 30;

9 APSNDGAAGLVPEAN--NETMALEP-----VAGASIAAPLTGQNNI 48

394 TISEKNGSKLGIGVATDVIIPIGPDGWPTTIAADKLIIPAGDYSITTEGNDIKTAQAYDT 453
325 -----TDFSGKLFVLSQRDHNDACRSHDAVIAT-----NSAKFTPKLGAIOIGTWEE 373
454 AAVVKNNTNFRGMYICGSLQRAWGDKKISNTAFITATRDGNEIKPSNTIDMTKLAVYQ- 512
374 DDVHINOPTK-----FTPVGLFENEGFNQWTLNPNYSALTNMGLAPPVAPFPFG 423
513 -DTHVEQEVOQTSDDTLALLGYTGIG-EAAGSNRRDRVRIS-----VLPAGARGG 561
424 EQILFRSHIPL 435
562 NHPFVKNISKL 573

RESULT 11
JQ2354
capsid protein - feline calicivirus (strain NADC)
C:Species: feline calicivirus
C>Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: JQ2354
R;Seal, B.S.; Kipach, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable
A:Reference number: JQ2354; MUID:94065683; PMID:7504075
A:Accession: JQ2354
A:Molecule type: mRNA
A:Residues: 1-668 <SEA>
A:Cross-references: GB:L09718; NID:g305104; PIDN:AAA16485.1; PID:g305105
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein

Query Match 9.8%; Score 277.5; DB 2; Length 668;
Best Local Similarity 32.1%; Pred. No. 1.6e-12;
Matches 79; Conservative 35; Mismatches 93; Indels 39; Gaps 9;

30 EPVAGASIAAPLTGQNNIIDPW-IRLNFVQAPNGEFTVSPRNSPGVLLNLELGPPLNPY 88
145 EPSQMTAADMATGKSVDSWEAFPSFTSVNW-----STSETQGKILFKQSLGPLNPY 200
89 LAHLRMYNGVAGGVEVQVLLAGNAFTAGKLVFAAAPPHPFLENISPGQIT---MPPHVI 145
201 LEHLKLVVWGSVGVFSGVFGGKLAIVVPP-----GVDPVQSTMLQYPHVL 255
146 IDVRLTEPVLPLDPVRNFFHYNQNEPRMLVAMLYTFLPILTIGELTNSRFP-----VPIDELYTSPNE 199
256 FDAQVDPVIFSIPLRSTLYHLMPTD-TTSLVIMVYNDLINPYNANDSSGCIV---- 310
200 CRVLTRSPDPDFNLYVPPTLESKTKPFTLPILTIGELTNSRFP-----VPIDELYTSPNE 255
311 -TVETKPGDPFKFHLKPPG-----SMLTHGSPDLPKSSSLMIGNRWYSDIT 359
256 SLVQOP 261
360 DFVRP 365

RESULT 12
VCWVFP9
coat protein - feline calicivirus (strain F9)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C>Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 28-Jul-2000
C:Accession: B43382; C45538; PQ0407; S23702
R;Carter, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.; Turner, P.C.
Virology 190, 443-448, 1992
A:Title: The complete nucleotide sequence of a feline calicivirus.
A:Reference number: A43382; MUID:92410623; PMID:1529544
A:Accession: B43382
A:Molecule type: Genomic RNA
A:Residues: 1-671 <CARL>
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879
R;Carter, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.

Arch. Virol. 122, 223-235, 1992
A:Title: Identification and sequence determination of the capsid protein gene of feline
A:Reference number: A45538; MUID:92117861; PMID:1731695
A:Accession: C45538
A:Molecule type: genomic RNA; protein
A:Residues: 1-671 <CAR>
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879
A:Experimental source: strain F9
A>Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)
R;Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.
J. Gen. Virol. 73, 2429-2433, 1992
A:Title: The cloning, sequencing and expression of a major antigenic region from the fel
A:Reference number: PQ0407; MUID:93019069; PMID:1402818
A:Accession: PQ0407
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 217-266, 'SI', 269-336, 'V', 338-395, 'S', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>
A:Cross-references: PIDN:AAB2353.1; PID:g257083
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F;177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.8%; Score 277.5; DB 1; Length 671;
Best Local Similarity 31.4%; Pred. No. 1.6e-12;
Matches 76; Conservative 38; Mismatches 97; Indels 31; Gaps 9;

30 EPVAGASIAAPLTGQNNIIDPW-IRLNFVQAPNGEFTVSPRNSPGVLLNLELGPPLNPY 88
145 EPSQMTAADMATGKSVDSWEAFPSFTSVNW-----STSETQGKILFKQSLGPLNPY 200
89 LAHLRMYNGVAGGVEVQVLLAGNAFTAGKLVFAAAPPHPFLENISPGQIT---MPPHVI 145
201 LEHLKLVVWGSVGVFSGVFGGKLAIVVPP-----GVDPVQSTMLQYPHVL 255
146 IDVRLTEPVLPLDPVRNFFHYNQNEPRMLVAMLYTFLPILTIGELTNSRFP-----VPIDELYTSPNESLV 203
256 FDAQVDPVIFSIPLRSTLYHLMPTD-TTSLVIMVYNDLINPYNANDSSGCIVTV 313
204 TRPSDPDPDFNLYVPPTLESKTKPFTLPILTIGELTNSRFP-----VPIDELYTSPNESLV 259
314 TKPGDPDFKHLKPPG-----SMLTHGSPDLPKSSSLMIGNRWYSDITDFVI 363
260 QP 261
364 RP 365

RESULT 13
C48562
coat protein - San Miguel sea lion virus (serotype 4)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: C48562
R;Neill, J.D.
Virus Res. 24, 211-222, 1992
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel
eins.
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Accession: C48562
A:Molecule type: genomic RNA
A:Residues: 1-703 <NEI>
A:Cross-references: GB:M87482; NID:g334886; PIDN:AAA16220.1; PID:g334888
A>Note: sequence extracted from NCBI backbone (NCBIP:113567)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F;89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.6%; Score 273.5; DB 1; Length 703;
Best Local Similarity 28.8%; Pred. No. 3.3e-12;
Matches 97; Conservative 51; Mismatches 150; Indels 39; Gaps 15;

8 AAPSNDAAGLVPEANNETM----ALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAPNGE 63


```
Db 151 ASDGPGSARI VTEBOGT VVQQPAPAPATATATATASTGK-SVDEQWMTFFSYHTSINW 209
QY 64 FTVSPRNSPGEVLLNLELGPENLPYLAHL SRMYNGVAGVEVQVLLAGNAFTAGKLVFAA 123
Db 210 STV---ESQKILYSQALNPSINPYLDHIAKLYSTWSGGIDVRFTVSGSGVFGKLAALL 266
QY 124 VPPHP-PLENISPGQITMPEHVIDVRLPVLPLPDVRNNFFHYNQNEPRMRLVAML 182
Db 267 VPPGVEPIESVS---MLQYPHVLFDAQOTEPVIFTIPDIRKTLFHSMDTD-TTKLVINP 322
QY 183 YTPLRNSGGDDVFTVSCRVLRPSDFDNVLPPTLESKTKPFTLPILTIGELTNSRF 242
Db 323 Y---ENGV-ENKTTCSIVETRPSADFTFALLKPP--GSLIKHGSIP-----SDLLPRNS 371
QY 243 PVPIDELYTSNPNESLVQVQ---NGRCALDGELOQ--TTQLLPTAICSFGRINQKVSQ 296
Db 372 AHWMGNRWSTISGFSVQVRVQSRNHRPDPDSTTTGWSPTPYVPIEI-----KIQKVG 426
QY 297 EN---HVMNQVTNINGTPTDPTGDPAPLGTDPDFSG 330
Db 427 NKKWFHVIDTKALVPGIP-DGWPDTTIPDETATNG 462
```

RESULT 14

```
VCMWFF
coat protein - feline calicivirus (strain CFI/68 FIV)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C>Date: 30-Jun-1992 #sequence revision 30-Jun-1992 #text_change 20-Aug-1999
C:Accession: A40507; B40507; T09246
R:Neill, J.D.; Reardon, I.M.; Heinrikson, R.L.
J. Virol. 65, 5440-5447, 1991
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline calicivirus
A:Reference number: A40507; MUID:91374597; PMID:1716692
A:Accession: A40507
A:Molecule type: Genomic RNA
A:Residues: 1-668 <NE1>
A:Cross-references: GB:132819; NID:G323874; PIDN:AAA42925.1; PID:G323875
A:Accession: B40507
A:Molecule type: protein
A:Residues: 373-379; 403-419; 481-489; 560-566 <NE2>
R:Neill, J.D.
submitted to the EMBL Data Library, April 1998
A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.
A:Reference number: Z16626
A:Accession: T09246
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: Genomic RNA
A:Residues: 1-668 <NE3>
A:Cross-references: EMBL:U13992; NID:G3056875; PIDN:AAC13993.1; PID:G537256
A:Experimental source: strain CFI/68 FIV
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted
```

```
Query Match 9.6%; Score 272.5; DB 1; Length 668;
Best Local Similarity 31.3%; Pred. No. 3.6e-12;
Matches 77; Conservative 37; Mismatches 93; Indels 39; Gaps 9;

QY 30 EPVAGASIAAPTGTGQNNIIDPW-IRLNFVQAPNGEFTVSPRNSPGEVLLNLELGPENLPY 88
Db 145 EPNAQWSTAADMATGKVDSEWSEAFPSHTSVNM---STSETQKILFKQSLGILLNPFY 200
QY 89 LAHLRMYNGYAGGVEVQVLLAGNAFTAGKLVFAAVPPHPPLENISPGQIT---MPPHVI 145
Db 201 LTHLAKLYVWAGSGVDVRSISGSGVFGKLAIVVPP-----GIDPVQSTSMLOYPHVL 255
QY 146 IDVRLTEPVLPLPDVRNNFFHYNQNEPRMRLVAMLYTPL-----RSNGSGDDVFTVS 199
Db 256 FPARQVEPVIFSIDRLSTLYLHMSDTD-TTSLVMVYNDLINPYANDNSNGGCIV---- 310
QY 200 CRVLTRPSDFDNVLPPTLESKTKPFTLPILTIGELTNSRFP-----VPIDELYTSNPN 255
```

```
Db 311 -TVEKPGDFKPHLLKPPG-----SMLTHGSIPSDLIKSSSLWIGNRFPWSDIT 359
QY 256 SLVQVP 261
Db 360 DFVIRP 365

RESULT 15
JQ2356
capsid protein - feline calicivirus (strain KCD)
C:Species: feline calicivirus
C>Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: JQ2356
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable regions
A:Reference number: JQ2354; MUID:94065683; PMID:7504075
A:Accession: JQ2356
A:Molecule type: mRNA
A:Residues: 1-668 <SEA>
A:Cross-references: GB:L09719; NID:G305107; PIDN:AAA16487.1; PID:G305108
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein

Query Match 9.4%; Score 266; DB 2; Length 668;
Best Local Similarity 29.6%; Pred. No. 1.1e-11;
Matches 79; Conservative 43; Mismatches 103; Indels 42; Gaps 11;

QY 11 SNDGAAGLVPEANNET--MALEPVAGASIAAPLTGQNNIIDPW-IRLNFVQAPNGEFTVS 67
Db 125 ADDGSI-VTPEQGTWVGVIAPNAQMTAADMATGKSVDSWEAFPSHTSVNM----S 179
QY 68 PRNSPGEVLLNLELGPENLPYLAHL SRMYNGVAGVEVQVLLAGNAFTAGKLVFAAVPPH 127
Db 180 TSETQKILFKQSLGILLNPFYLAHLAKLYVWAGSGIEVRSISGSGVFGKLAIVVPP- 238
QY 128 PPLENISPGQIT---MPPHVIDVRLTEPVLPLPDVRNNFFHYNQNEPRMRLVAMLYT 184
Db 239 ----GVDPVQSTSMLOYPHVLFDARQVEPVIFSIDRLSTLYLHMSDTD-TTSLVMVY 293
QY 185 PL-----RSNGSGDDVFTVSCRVLRPSDFDNVLPPTLESKTKPFTLPILTIGELT 238
Db 294 DLINPYANDNSNGGCIV-----TVETKFGSDFRPHLLKPPG-----SVLTHGSVP 338
QY 239 NSRFP-----VPIDELYTSNPNESLVQVP 261
Db 339 SDLIPTSSSLWIGNRYWTNITNFI 365
```

Search completed: June 1, 2004, 13:55:28
Job time : 11.323 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:36:02 ; Search time 10.4581 Seconds
(without alignments)
4985.230 Million cell updates/sec

Title: US-09-926-799-9
Perfect score: 2863
Sequence: 1 MKMASNDAPSDNGAASLVP.....VNQFYSLAPMTGTGNGRRRIQ 542

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1980	69.2	542	2 S60616	capsid protein - h
2	1833.5	64.0	539	2 S40111	capsid protein - h
3	1179	41.2	530	2 B37471	capsid protein - N
4	1159.5	40.5	546	2 B37491	major capsid prote
5	340	11.9	2344	2 S55399	genome polyprotein
6	339	11.8	2344	1 RRWRH	genome polyprotein
7	339	11.8	2344	2 S64740	genome polyprotein
8	326.5	11.4	576	2 A53982	capsid protein - E
9	290	10.1	668	2 JQ2354	capsid protein - f
10	279	9.7	668	1 VCNWFF	coat protein - fel
11	276.5	9.7	671	1 VCNWFC	coat protein - fel
12	271.5	9.5	668	1 VCNWFC	coat protein - fel
13	269.5	9.4	702	1 A48562	coat protein - San
14	268.5	9.4	668	2 JQ2356	capsid protein - f
15	240	8.4	703	1 C48562	coat protein - San
16	145.5	5.1	2333	1 GNNY2F	genome polyprotein
17	130	4.5	2332	1 GNNY4F	genome polyprotein
18	123.5	4.3	746	2 T46821	siderophore recept
19	123.5	4.3	746	2 A95420	RhtA Rhizobactin r
20	120.5	4.2	1742	2 S76110	hypothetical prote
21	119	4.2	701	2 A11501	internalin protein
22	118	4.1	2194	1 GNNY7E	genome polyprotein
23	116	4.1	545	2 S28117	gas-vesicle operon
24	115.5	4.0	3085	2 T00327	polyprotein - infe
25	114.5	4.0	785	2 S75819	general secretion
26	114	4.0	2336	2 S37077	genome polyprotein
27	111	3.9	940	2 D89723	protein F39D8.1b [
28	111	3.9	945	2 T21998	hypothetical prote
29	111	3.9	13055	2 T16580	hypothetical prote

30 110.5 3.9 1165 2 F90877 probable host spec
31 110.5 3.9 1165 2 D85842 probable tail fibre
32 110.5 3.9 5376 2 T42215 zonadhesin - mouse
33 110 3.8 1196 2 S35994 DNA repair protein
34 110 3.8 2290 1 GNNYE genome polyprotein
35 109.5 3.8 852 1 GNNYHA genome polyprotein
36 109.5 3.8 1109 2 T18536 receptor-like prot
37 109.5 3.8 2468 2 A83412 hypothetical prote
38 109 3.8 811 2 T00036 capsid protein pre
39 109 3.8 2206 2 S03822 genome polyprotein
40 108.5 3.8 900 1 GNNYV genome polyprotein
41 108.5 3.8 1358 2 A03905 genome polyprotein
42 108.5 3.8 2227 1 GNNYHM genome polyprotein
43 108.5 3.8 2227 1 GNNYHR genome polyprotein
44 108.5 3.8 2227 1 GNNYMK genome polyprotein
45 108.5 3.8 2227 1 GNNYHB genome polyprotein

ALIGNMENTS

RESULT 1

S60616
capsid protein - human calicivirus (strain Melksham)

C:Species: human calicivirus

A:Variety: strain Melksham

C:Date: 23-May-1997 #sequence_revision 23-May-1997 #text_change 28-Jul-2000

C:Accession: S60616

R:Green, S.M.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.

Varian Research, 1997

A:Title: Capsid diversity in small round-structured viruses: molecular characterization

A:Reference number: S60615; MUID:96136658; PMID:8533462

A:Accession: S60616

A:Molecule type: genomic RNA

A:Residues: 1-542 <GRS>

A:Cross-references: EMBL:X81879; NID:G976077; PIDN:CAA57462.1; PID:G976079

A:Experimental source: strain Melksham

A:Note: it is uncertain whether Met-1 or Met-3 is the Initiator

C:Superfamily: human calicivirus capsid protein

C:Keywords: capsid protein; coat protein

Query Match 69.2%; Score 1980; DB 2; Length 542;

Best Local Similarity 67.6%; Pred. No. 1.2e-139;

Matches 369; Conservative 81; Mismatches 88; Indels 8; Gaps 5;

QY 1 MKMASNDAPSDNGAASLVPESINMETPLEPVAGASIAAPVAGQTNIIDPWIRTNFVQAP 60

DB 1 MKMASNDAPSDTGGAGLVPSNNEVMALEPVAGAGLAAPVTGQTNIIDPWIRANFVQAP 60

QY 61 NGEFTVSPRNGSGEILLNLELGPDLNPLYLAHLSRMVYAGGVGVQLLAGNAFTAGKTL 120

DB 61 NGEFTVSPRNPAGEVLLNLELGPDLNPLYLAHLSRMVYAGGVGVQLLAGNAFTAGKLV 120

QY 121 FAATPPNLFVDMISPAQITMLPHLIVDRTLEPIMPLPDVNRVYHFNQPRMLVA 180

DB 121 FAAPPPHFPVENLSPKITMFPFHVIDVRLTLEPVLPLPDVNRNPFPHYKQDPKMRIVA 180

QY 181 MLYTLRNSGDDVFTVSCVRLTRPTDFEPIYLVPPSVESKTKPFTLPILTISELTS 240

DB 181 MLYTLRNSGDDVFTVSCVRLTRPTDFEPIYLVPPSVESKTKPFTLPILTISELTS 240

QY 241 RFPPIEQLYTAPNETNVVQCGRCTLDELGGTTLQLSSAVCFLOGRVTADNGDNWDQ 300

DB 241 RFPVIDQMTSPNEVISVQCGRCTLDELGGTTLQLQVSGICAFKGEVTAHLHDNDHL 300

QY 301 NLLQUTYNGASDYDTDEVPAPLQDFSGMLYGVLTQDNVNVSTGEAKNAK---GIYIS 357

DB 301 NNVVITNLNGSPFDPSEDIPALGVDPFGQGVFGVISQDKQNAAGHSEPANRGHDAVVP 360

QY 358 TTSGRFTPKIGSIGHLS-ITEHVHFNQQRFTFVGVAVDENTPFQOQVLPVHAGSLANT 416

DB 361 TYTAQYTKLGQIQIGTQWTDLTVNQPVKFTFVG--LNDTEHFNQMVVPRYAGALNLT 418

QY	417	NLAPAVAPTFEGEQLLFFRSVPVCVQGLQGQDAFIDCLLPQEWNVHFYQEAAPSQADVAL	476
Db	419	NLAQSVAPVFFGERLLFFRSHLP-LKGGYGNPA-IDCLLPQEWVQHFYQEAAPSMSEVAL	476
QY	477	IRYVNPDTGRTLFEAKLHRSFGFIIVSHTGAYPLVVPNGHFRDSDWVNQFYSILAPMGTCN	536
Db	477	VRYINPDTRGALFEAKLHRAFGMTVSSNTSAPVVVPVANGYFRDSDWVNQFYSILAPMGTCN	536
QY	537	GRRRQ 542	
Db	537	GRRRVQ 542	
RESULT 2			
S40111			
capsid protein - human calicivirus (strain Bristol isolate B493)			
C/Species: human calicivirus			
A/Variety: strain Bristol isolate B493			
C/Date: 25-Dec-1994 #sequence_revision 27-Feb-1997 #text_change 28-Jul-2000			
C/Accession: S40111			
R/Green, S.M.; Dingle, K.E.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, submitted to the EMBL Data Library, December 1993			
A/description: Human enteric caliciviridae: a new prevalent SRSV group defined			
A/Reference number: S40111			
A/Accession: S40111			
A/Molecule type: genomic RNA			
A/Residues: 1-539 <GRE>			
A/Cross-references: EMBL:X76716; NID:G436410; PIDN:CAA54134.1; PID:G436411			
A/Experimental source: human enteric calicivirus strain Bristol isolate B493			
C/Superfamily: human calicivirus capsid protein			
C/Keywords: capsid protein; coat protein			
Query Match 64.0%; Score 1833.5; DB 2; Length 539;			
Best Local Similarity 64.0%; Pred. No. 9.6e-129;			
Matches 349; Conservative 74; Mismatches 109; Indels 13; Gaps 6			
QY	1	MKQASNDAPSDGASLYPEGINETWPLEPVAGASIAAPVAGQTNLIIDPIRTNFVQAP	60

C:Accession: B37491
R:Amenden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
Science 259, 516-519, 1993
A:Title: Sequence and genome organization of a human small round-structured (Norwalk-like
A:Reference number: A37491; MUID:93142023; PMID:8380940
A:Accession: B37491
A:Status: not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-546 <LAM>
A:Cross-references: GB:L07418; NID:g1236787; PIDN:AAA2984.1; PID:g295114
A:Note: sequence extracted from NCBI backbone (NCBIP:123458)
A:Note: small round-structured virus, SRSV, Norwalk virus, Norwalk-like virus, serotype
C:Superfamily: human calicivirus capsid protein
C:Keywords: glycoprotein
F:303,340,441/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.5%; Score 1159.5; DB 2; Length 546;
Best Local Similarity 45.3%; Pred. No. 1.6e-78;
Matches 258; Conservative 84; Mismatches 176; Indels 51; Gaps 17;
QY 1 MOWASNDAAPSNDGAA---SLVPEGIN--ETMPLFPVAGASTAAPVAGQTNIIDPWRTN 55
DB 1 MOWASKDAPQSDAGSAGAQVPE-VNTADPLPMEPVAGPTTAVATAGQVNMIDPWVNN 59
QY 56 FVOAPNGEFTVSPRNSPGEILLNLGLPDLNPLYLHLRMYNGYAGGVEQVLLAGNAPT 115
DB 60 FVQSQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGVGNRVRILLAGNAPS 119
QY 116 AKGILFAAPLPPNPLVDMISPAQITMLPHLIVDVRLTPIMLPDRNVFHFENNOPOPR 175
DB 120 AKGIIVCCVPPGFTSSSLTIAQTAFPHVIAVRLTEPLEMDRNVLYHTNON-QPT 178
QY 176 MELVAMLYTLPLRS-NGSGD-DVFTVSCRVLTPDFFFIYLVPPSVESKTKPFTLPILT 233
DB 179 MELVCMLYTLPLTGGSGNSDSFVAGRVLTAPSSDSFSLFLVPTTISQKTAFTVPNIP 238
QY 234 ISELNRPPIPIEOLYTAPNETNVVQCONGRCTLDGELQGTQLLSAVCFLOQRTVAD 293
DB 239 LQTLNSRPPSIQGMILSPDASQVVFQNGRLDGLLGTTPATSGOLFVRGKI---295
QY 294 NGDNWDNLQLTYPNGASYDPTDEVPAPLQGFQDSGLMYGLVLTQDNVNVSTGEAKNAG 353
DB 296 ---NQARTLNUTEVDGKPFMAFDS-PAPVGPDPFGKCDWHMRISKTPN-NTGSGDPMRS 350
QY 354 IYISTTSKFTPKIGSIGLHSITEHVHP-----NQSRFTTPVGVAVDENTPFOQ 402
DB 351 VSVQTNVQGFVPHLGSIQDEVEN--HPTGDIYGTIEWISQPS--TPPG-----TDINL 400
QY 403 WVLPHYAGSLALNTNLAPAVAPTFPGEQLLFRRSRVPCVQGLQGQDA--FIDCLLPQEWV 460
DB 401 WEIPDYGSSLSQAANLAPVPFPGFGEALVYFVSAPP---GPNRSAPNDVPCLLPQBYI 457
QY 461 NHFYQEAAPSQADVALIRVYVNDTGTLEFEAKLHRSGLFTV--SHTGAYPLVVPNGHPR 518
DB 458 THFVSEQAPTMDGALLHYVDPTNRNLGEFKLXPGGYLTCPVNGVGAGPQOQLPLNGVPL 517
QY 519 FDSWYNQFSLAPMGTN-----GRRRI 541
DB 518 FVSWVSRFYQLAPVGTASTARGELGVRI 546

RESULT 5

S55399
genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)
C:Species: rabbit hemorrhagic disease virus
A:Variety: isolate BS89
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 18-Jun-1999
C:Accession: S55399
R:Rossi, C.
submitted to the EMBL Data Library, May 1995
A:Reference number: S55399
A:Accession: S55399
A:Molecule type: genomic RNA

A:Residues: 1-2344 <ROS>
C:Cross-references: EMBL:X87607; NID:g854640; PIDN:CAA60910.1; PID:g854641
A:Experimental source: isolate BS89
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: polyprotein

Query Match 11.9%; Score 340; DB 2; Length 2344;
Best Local Similarity 26.4%; Pred. No. 1.7e-16;
Matches 141; Conservative 70; Mismatches 209; Indels 114; Gaps 22;
QY 19 VPEGINTEMPLPVPVAGASIAAPVAG-QTNIIDPMI-----52
DB 1765 VMEGKARTAPQGEAACTATTATSVPTTGDGLDPGVVATTSVTAENSSASIAATAGIGGPP 1824
QY 53 -----RTNFVOAPNGEFTVSPRNSPGEILLNLGLPDLNPLYLHLRMYNGYAGGV 103
DB 1825 QVDDQOQETWRTNFYY--NDVFTWSVADAPGSILYTVQHSPPQNNPFTAVLSQMYAGWAGM 1882
QY 104 EVOVLLAGNATAGKILFAATPPNPLVDMISPAQITMLPHLIVDVRLTPIMLPDRNV 163
DB 1883 QFRFTVAGSGVFGGRLVAIVPPG--IEIGPLEVRQPFHVVIDARSLEPVTITPDLRP 1940
QY 164 VFYHENNOPOPRMRLVAMLYTLPLRSNGSGDDVFTVSCRVLTPDFFFIYLVPPSVESK 223
DB 1941 NMYHTGDPGLVTLVLVSYNNL-INPPGSGTSAIQVTETRPSEDFEFVMTAIFS--SK 1997
QY 224 TKPFTLP--ILTISELT-----NSRPPPIEOLYTAPNETNVVQCONGRCTLDGELQGT 277
DB 1998 TVDSISPAGLLTPVLTGVGNDNRWNGQIVGLQPVGGFST--C-NRHNWLNSTGVWSS 2054
QY 278 LLSSAVCFLOQRTVADNGDNWDNLQLTYPNGASY--DPTDEVPAPIGTQD-----FSG 330
DB 2055 PRFADIDHRRG-SASYPGSN-ATNVLQFWYANAGSAVDNPISQV-APDGFDPMSFVPFG 2111
QY 331 -----MLGVLTQDN---VNVSTGRAKN-----AKGIYI 356
DB 2112 PCIPAGWVGCAIWNSSGAPNVTVQAYELGFATGAPGNLQPTNTSGAOTVAKSIVA 2171
QY 357 STT-----SGKFTPKIGSIGLHSITEHVHPNQSRF-----TPVGVAVDENTPFOQWYL 405
DB 2172 VVTGTAQNPAGLFVWASGVISTPNANAITPDPDRIVTTPCTPAAAAPVGNKTPIM---2227
QY 406 PHYAGSLALNTNLAPAVAPTFPGEQLLFRRSRVPCVQGLQGQDAFIDCLLPQEW 459
DB 2228 ---FASVVRRTGDVNVATAGSANGTQGTGSOPLPVTIGL-SLNYSYSSALMPQOF 2277

RESULT 6

RRWRH
genome polyprotein - rabbit hemorrhagic disease virus
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C:Species: rabbit hemorrhagic disease virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 29-May-1998
C:Accession: A41039
R:Meyers, G.; Wirblich, C.; Thiel, H.J.
Virology 184, 664-676, 1991
A:Title: Rabbit hemorrhagic disease virus--molecular cloning and nucleotide sequencing
A:Reference number: A41039; MUID:91361557; PMID:1840711
A:Accession: A41039
A:Molecule type: genomic RNA
A:Residues: 1-2344 <MSY>
C:Cross-references: GB:M67473
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: nucleotidyltransferase

Query Match 11.8%; Score 339; DB 1; Length 2344;
Best Local Similarity 27.5%; Pred. No. 2e-16;
Matches 145; Conservative 71; Mismatches 214; Indels 98; Gaps 25;
QY 8 AAPSNDGAASL-----VPEGINTEMPLPVA-----GASIA-APVAGQTNIID---49
DB 1772 AAPQGEAACTATTASVPTTGDMDPGVVATTSVTAENSSASIAATAGIGGPPQVDOOE 1831

50 PHIRNFVQANGTEFTSPRSPGELLNLELGLPDLNPLYLALHLSRMVINGAGGVEQVLL 109
 1832 TW-RNFYY--NDVFTWSVADAPGILTYVQHSQNNPFTAVLSQMYAGWAGGQFRFIV 1888
 110 AGNAFTAGKILFAAIPPNFLVDMISPAQITMLPHLIVDVRLEPIMPLDVRNVFFHN 169
 1889 AGSGVFGRLVRAVIPP--IEIGPGLVVRQFPFHVVVDARSLEPVTITMPDLRPNMYHT 1946
 170 NOPQPMRLVAMLYTPLSNGSGDDVFTVSCRVLTRPTDPEFIYLVPPSVESKPKFTL 229
 1947 GDFGLVPLTVLSVYNNL--INPFGSTSAIQVTVETRPSEDFEFWIRAPS--SKTVDSIS 2003
 230 P--ILITISELT---NSRFPPIEQLYTAPNETNVVQCNGRCTLDGELGTQTQLSSAV 283
 2004 PAGLLTTPVLTGVGDNRWNGQIVGLQPVPGFST--C-NRHWNLNGSTYGWSSPFRGDI 2060
 284 CFLQRTVADNGDNQNLQTYPNGASY--DPTDEVPAPLGTQD-----FSG----- 330
 2061. DHRRG--SASYSGSN-ATNVLFQWYANAGSAIDNPISQV-APDGFPMDFVFPNGGIPAA 2117
 331 --MLYGLVLTQDN---VNVSTGEAKN-----AKGIYISTT--- 359
 2118 GWVFGAIWNSGAPNVTVOAYELGPATGAPGNLQPTNTSGAQTVAKSIYAVVTGTA 2177
 360 ---SGKFTPKIGSIGLHSITEHVHPNQQRFP-----TPGVAVDENTPQQWVLPHYAGS 411
 2178 QNPAGLFVWASGIISTENAGRIITYTPQDRIVTTPGTAAAPVGNKTPIM-----FAS 2230
 412 LALNTNLAPAVATTFGEQLLFRSRVPCVQGLQGDADFDCLLPOEW 459
 2231 VVRRTGDVNTAGSANGTQYGTGSQLPVTIGL-SUNNYSSALMPCQF 2277

RESULT 7
 S64740 genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)
 N;Contains: VP60 protein
 C;Species: rabbit hemorrhagic disease virus
 A;Variety: isolate AST/89
 C;Date: 12-Jul-1996 #sequence revision 26-Jul-1996 #text_change 20-Jun-2000
 C;Accession: S64740; S46944; S49018; S65012
 R;Boga, J.; Casais, R.; Martin-Alonso, J.; Boga, J.; Parra, F.
 submitted to the EMBL Data Library, May 1995
 A;Description: Genomic organization of rabbit hemorrhagic disease virus determined by di
 A;Reference number: S64740
 A;Accession: S64740
 A;Molecule type: genomic RNA
 A;Residues: 1-2344 <CAS>
 A;Cross-references: EMBL:Z49271; NID:gl182032; PIDN:CAA89265.1; PID:gl150552
 A;Experimental source: isolate AST/89
 R;Boga, J.; Casais, R.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra, F.
 submitted to the EMBL Data Library, July 1993
 A;Description: Molecular cloning, sequence and expression of the capsid protein gene f
 A;Reference number: S46944
 A;Accession: S46944
 A;Molecule type: genomic RNA
 A;Residues: 1650-2344 <BOG>
 A;Cross-references: EMBL:Z24757; NID:g515622; PIDN:CAAB08081.1; PID:g515623
 A;Experimental source: isolate AST/89
 R;Parra, F.; Boga, J.A.; Marin, M.S.; Casais, R.
 Virus Res. 27, 219-228, 1993
 A;Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus supp
 A;Reference number: S49018; MUID:93255896; PMID:8488721
 A;Accession: S49018
 A;Molecule type: genomic RNA
 A;Residues: 1650-1796 <PAR>
 A;Cross-references: EMBL:Z24757
 A;Experimental source: isolate AST/89
 A;Accession: S65012
 A;Molecule type: protein
 A;Residues: 1767-1779;1875-1877,'X',1879-1881,1936-1938,'X',1940-1941 <PAM>
 C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase
 C;Keywords: blocked amino end; polyprotein

Query Match 11.8%; Score 339; DB 2; Length 2344;
 Best Local Similarity 26.9%; Fred. No. 2e-16;
 Matches 144; Conservative 61; Mismatches 213; Indels 118; Gaps 20;

19 VPEGINETMPLFPVAGASIAAPVAG-QTNIIIDPMI----- 52
 1765 VMEGKARTAPOGEAAGTATTASVPGTTTGDMDGPGVAVTTVVVTAENSSASIIATAGIGOPP 1824
 53 -----RTNFPVQAPNGEFTVSPRSPGELLNLELGLPDLNPLYLALHLSRMVINGAGG 103
 1825 QQVDQOETWRITNFYY--NDVFTWSVADAPGSLTYVQHSQNNPFTAVLSQMYAGWAGGM 1882
 104 EVQVTLAGNAFTAGKILFAAIPPNFLVDMISPAQITMLPHLIVDVRLEPIMPLDVRNV 163
 1883 QRFIVAGIVGFGRLVRAVIPP--IEIGPGLVVRQFPFHVVVDARSLEPVTITMPDLRP 1940
 164 VFYHFNQPMRLVAMLYTPLSNGSGDDVFTVSCRVLTRPTDPEFIYLVPPSVESK 223
 1941 NMYHPTGDPLVPLTVLSVYNNL--INPFGSTSAIQVTVETRPSEDFEFWIRAPS--SK 1997
 224 TKPFTLP--ILITISELT---NSRFPPIEQLYTAPNETNVVQCNGRCTLDGELGTQT 277
 1998 TVDSISPAGLTTPVLTGVGDNRWNGQIVGLQPVPGFST-----CNRHWNLNGSTY 2050
 278 LLSAVCFLOQ--RTVADNGDNQNLQTYPNGASY--DPTDEVPAPLGTQD-----F 328
 2051 GWSPPRGDIHRRGSASYPGNATNVLFQWYANAGSAIDNPISQV-APDGFPMDFVFP 2109
 329 SG-----MLYGLVLTQDN---VNVST-----GEAKNAKGI 354
 2110 NGCPGIPAAGWVFGAIWNSGAPNVTVOAYELGPATGAPGNLQPTNTSGAQTVAKSI 2169
 355 YISTT-----SGKFTPKIGSIGLHSITEHVHPNQQRFP-----TPGVAVDENTPQQW 403
 2170 YAVVTGTAQNPAGLFVWASGVISTPSANAITYPQDRIVTTPGTAAAPVGNKTPIM-- 2227
 404 VLPHYAGSLALNTNLAPAVATTFGEQLLFRSRVPCVQGLQGDADFDCLLPOEW 459
 2228 ----FASVVRRTGDVNTAGSANGTQYGTGSQLPVTIGL-SUNNYSSALMPCQF 2277

RESULT 8
 A53982 capsid protein - European brown hare syndrome virus
 C;Species: European brown hare syndrome virus
 C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 28-Jul-2000
 C;Accession: A53982
 R;Wirdlich, C.; Meyers, G.; Ohlinger, V.F.; Capucci, L.; Eskens, U.; Haas, B.; Thiel, H.
 J. Virol. 68, 5164-5173, 1994
 A;Title: European brown hare syndrome virus: relationship to rabbit hemorrhagic disease
 A;Reference number: A53982; MUID:94309183; PMID:7518531
 A;Accession: A53982
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-576 <NIR>
 A;Cross-references: GB:U09199
 C;Superfamily: human calicivirus capsid protein

Query Match 11.4%; Score 326.5; DB 2; Length 576;
 Best Local Similarity 26.9%; Pred. No. 2.1e-16;
 Matches 127; Conservative 67; Mismatches 157; Indels 121; Gaps 24;

3 MASNDAAPSNDGAASLVPEGINETMPLFPVAGASIAAPVAGQTNIIIDPWIRTNFVQAPNG 62
 31 VASTDVVTADNVAAS-----VATAGTGGP-PQASQESW-RWNFFV--ND 72
 63 EFTVSPRSPGELLNLELGLPDLNPLYLALHLSRMVINGAGGVEQVLLAGNAFTAKILPA 122
 73 VFTWSVTDAPGSLTYVQHSQNNPFTAVLSQMYAGWAGGMQFRFIVAGSGIFGRLVCA 132
 123 AIPNPNFLVDMISPA-QITMLPHLIVDVRLEPIMPLDVRNVFVHFNNQPQPRMRLVAM 181

Db 133 IIPPGI---QIQGLEVRQPHVVIDARSLEPVTITMPDLRPMVHTPTGDPGLVPTLVVS 189
QY 182 LYTPLASNGDGDVFTVSCRVLTRPTDPDEFIYLPVPSVESKT----KP---FTLPILRI 234
Db 190 VYNNL-INFPGGTSAIQVTVETRPSEDEFVLIRAPS--SKTVDSVNSWLLTTPVLT- 245
QY 235 SELTNSRFPPIEQLYATARNVTVQCNGRCRTLDGELG----- 274
Db 246 GAGSDNRWGAIVGLQPVPGFST---SNRHNMNGSTYGSWSPRDDIDHPSGNVSYPT 302
QY 275 -----TTQLSS-----AVCFLLGRVTADN-----GDNDQD- 301
Db 303 GSATNTIETWYANAGTATTNPTISNIAPDGFDPDGAIPP-SGTITPGAWVGFQVWNASN 361
QY 302 -----LLQUTYNGA--SYDPTDEVPAPLGTQDFSMGLYGLVLTQDNVNVSTGEAKN 350
Db 362 GTPYVGTQVAYELGFANGAPSSIRP---VTTTGAQLVAKSIYGVIAIQNS-SAGIIFL 417
QY 351 AKGIYIST---TSGKETPKIGSLHSITEHVHPNQOSRFTTPVGVAVDENTP 399
Db 418 SKGM-VSTPGVAATTVPQSAI---VT-----TPGTPVAAPIGKNTP 456

RESULT 9
JQ2354
capsid protein - feline calicivirus (strain NADC)
C:Species: feline calicivirus
C>Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: JQ2354
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable
A:Reference number: JQ2354; MUID:94065683; PMID:7504075
A:Accession: JQ2354
A:Molecule type: mRNA
A:Residues: 1-668 <SEA>
A:Cross-references: GB:L09718; NID:G305104; PIDN:AAA16485.1; PID:G305105
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein

Query Match 10.1%; Score 290; DB 2; Length 668;
Best Local Similarity 23.4%; Pred. No. 1.4e-13;
Matches 138; Conservative 88; Mismatches 201; Indels 164; Gaps 30;
QY 11 SNDGAASLVPEG-INETMLEPVAGASIAAPVAGQTNIIIDPW-IRTNFVOAPNGEFTSP 68
Db 125 TDDGSITTEQGTGTVGVIAEBSAQMSAADMATGKSDVSEWEAPFSPHTSVNW----ST 180
QY 69 RNSPGEILLNLELGPDLNLYLAHLSRMVNGYAGGVEQVLLAGNAFTAGKILFAAIPPNF 128
Db 181 SETQKILFKQSLGPLLNPYLEHLSKLYVWAGSVSEVRFSGSGVFGGKLAIVVPPG- 239
QY 129 LVDMISPAQITMLPHLIVDVRTLEPTMPLPDVRNVFVHFNQPO-PRMRLVAMLYTPL- 186
Db 240 -VDPVQSTSMLOYPHVLFDARQVDFVIFSDPLRSTLYHL--MPDPTTSLVIMVYNDLI 296
QY 187 -----RNSGSGDDVFTVSCRVLTRPTDPDEFIYLVPP-----SVESK----- 223
Db 297 NPYANDSNSGCIV-----TVETKPGDPFKFLLKPPGSMLTGSPVSLIPKSSSLWIG 351
QY 224 -----TKPFTL-----PI-LTISELTNSRFPPIEQY 250
Db 352 NRYWSDITDFVVRFPFQANRHFNFQETAGWSAPRFRPITITISBSKSLGIGVATDY 411
QY 251 TAPN-----ETNVQCONGRCTLDGELQGTQLLSSAVCFLOGRTVADNGNDQNLQ 304
Db 412 IVPGIPDGPDPDTTIAE-----DLTPAGDYAIT-----SGNGD----- 444
QY 305 LTPVNGASYPDTPDEVAPLGTODFGSM-LYGVLTQDNVNVSTGEAKNAGIYISTT---S 360
Db 445 ---ITGSEYDSTEVK---NNTNFRGMVYCGSLQR-----AWGDKKISNTAFITTAKEG 494
QY 361 GKFTPK-----IGSIGLHSITEHVHPNQOSR-----FTPVG---VAVDENTPQQWYL 405

Db 495 NKIRPSTNTDMTKIAVYQDT-HVGBEVQTSDDALLALGYTGIGEOAIGSDRDRVVRISVL 553
QY 406 PHYAGSLALNTNLAPAVAPTFPGEQQLLFFRSRVPVCVQGLQGQDAFIDCULL---PQSWNVH 462
Db 554 PEVGAR-----CGNHPIFYKNSIKLGYVIRSIDVFNSQLHTSRQLSLNH 598
QY 463 FYOEAPASQADVALLRYNVPDTRTLFEAKLHRS--FITVSHTG--AYPL 509
Db 599 YLL-----SPDSPAVYRII--DSNGSWFOIDSEGFSFVGISISIGKLEYPL 643

RESULT 10
VCWVFF
coat protein - feline calicivirus (strain CFI/68 FIV)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Aug-1999
C:Accession: A40507; B40507; T09246
R:Neill, J.D.; Reardon, I.M.; Heinrichson, R.L.
J. Virol. 65, 5440-5447, 1991
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline calici
A:Reference number: A40507; MUID:91374597; PMID:1716692
A:Accession: A40507
A:Molecule type: genomic RNA
A:Residues: 1-668 <NEI>
A:Cross-references: GB:M32819; NID:G323874; PIDN:AAA42925.1; PID:G323875
A:Accession: B40507
A:Molecule type: protein
A:Residues: 373-379; 403-419; 481-489; 560-566 <NE2>
R:Neill, J.D.
submitted to the EMBL Data Library, April 1998
A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.
A:Reference number: Z16626
A:Accession: T09246
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: genomic RNA
A:Residues: 1-668 <NE3>
A:Cross-references: EMBL:U13992; NID:G3056875; PIDN:AAC13993.1; PID:G537256
A:Experimental source: strain CFI/68 FIV
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.7%; Score 279; DB 1; Length 668;
Best Local Similarity 23.2%; Pred. No. 9.2e-13;
Matches 137; Conservative 81; Mismatches 210; Indels 162; Gaps 28;
QY 11 SNDGAASLVPEGINETMP-----LEPVAGASIAAPVAGQTNIIIDPW-IRTNFVOAPNGEFT 65
Db 125 ADDGSITTPEQG---TMVGVIAEPNAQMSAADMATGKSDVSEWEAPFSPHTSVNW--- 178
QY 66 VSPRNSPGEILLNLELGPDLNLYLAHLSRMVNGYAGGVEQVLLAGNAFTAGKILFAAIP 125
Db 179 -STSETQKILFKQSLGPLLNPYLEHLSKLYVWAGSVSDVRFSGSGVFGGKLAIVVPP 237
QY 126 PNFLVDMISPAQITMLPHLIVDVRTLEPTMPLPDVRNVFVHFNQPO-PRMRLVAMLYT 185
Db 238 PG-IDPVQSTSMLOYPHVLFDARQVDFVIFSDPLRSTLYHLSMDTD-TTSLVIMVYND 294
QY 186 L-----RNSGSGDDVFTVSCRVLTRPTDPDEFIYLVPP-----SVESK----- 223
Db 295 LINPYANDSNSGCIV-----TVETKPGDPFKFLLKPPGSMLTGHSIPSLIPKSSSLW 349
QY 224 -----TKPFTL-----PI-LTISELTNSRFPPIEQ 248
Db 350 IGNRFWSDITDFVVRFPFQANRHFNFQETAGWSAPRFRPITITISVKSAGKIGVAT 409
QY 249 LYTAPN-----ETNVQCONGRCTLDGELQGTQLLSSAVCFLOGRTVADNGNDQNL 302
Db 410 DIVVIGIPDGPDPDTTIAE-----TIPGELVPV-----GDYAITNGTND--- 444
QY 303 LQLTYPNGASYPDTPDEVAPLGTODFGSM-LYGVLTQ-----DNNVNVSTGEAKNAG 353

```
Db 445 ----ITTAQYDAAEI---RNNTFRGWYICGSLQRAWGDKKISNTAFITTTGVDGAKL 497
QY 354 IYVSTSGKFTPKIGSIGLHSTEHVHNQOS-----RTPVG---VAVDENTPQQWVLP 406
Db 498 IPSNTID---QTKIAVFQDTHANKHVQSDDTLALLGYTGIGEEAIGADRVRVRSVLP 554
QY 407 HYAGSLALNTLAPAVATPFCEQLLFFRSRVPVCVQGLQGQDAFIDCLL---PQWVNHF 463
Db 555 ERGAR-----GGNHPIPHKNSIKLGVVIRSIDVFNSQLHTSRQLSLNHY 599
QY 464 YQEAAPSQADVALIRYVNDTQRTLFEAKLHRSQ--FTVSHTG--AYPL 509
Db 600 LL----SPDSFAVYRII--DSNGSWFDIGIDNDGFSFVGVSIGKLEPFL 643
```

RESULT 11

```
VCWMF9
coat protein - feline calicivirus (strain F9)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C>Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text_change 28-Jul-2000
C:Accession: B43382; C45538; FQ0407; S23702
R:Cartier, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.; P.C.
Virology 190, 443-448, 1992
A:Title: The complete nucleotide sequence of a feline calicivirus.
A:Reference number: A43382; MUID:92410623; PMID:1529544
A:Accession: B43382
A:Molecule type: genomic RNA
A:Residues: 1-671 <CAR1>
A:Cross-references: GB:M86379; NID:G323877; PIDN:AAAY9327.1; PID:G323879
R:Cartier, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.
Arch. Virol. 122, 223-235, 1992
A:Title: Identification and sequence determination of the capsid protein gene of feline
A:Reference number: A45538; MUID:92117861; PMID:1731695
A:Accession: C45538
A:Molecule type: genomic RNA; protein
A:Residues: 1-671 <CAR2>
A:Cross-references: GB:M86379; NID:G323877; PIDN:AAAY9327.1; PID:G323879
A:Experimental source: strain F9
A:Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBI:77462)
R:Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.
J. Gen. Virol. 73, 2429-2433, 1992
A:Title: The cloning, sequencing and expression of a major antigenic region from the fel
A:Reference number: FQ0407; MUID:93019069; PMID:1402818
A:Accession: FQ0407
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 217-266, 'SI', 269-336, 'V', 338-395, 'S', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>
A:Cross-references: PIDN:AA823553.1; PID:G257083
A:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted
```

```
Query Match 9.7%; Score 276.5; DB 1; Length 671;
Best Local Similarity 25.1%; Pred. No. 1.4e-12;
Matches 131; Conservative 68; Mismatches 171; Indels 151; Gaps 28;
```

```
QY 11 SNDGAASLPVGINETMP-----LEPVAGASTAAPVAGQNTIIDPW-IRTNFVQAPNGEFT 65
Db 125 ADDGSITAEQ---TWGVGIAEPSQAQMAADMATGKSDSEWEAFFSFHTSVNW--- 178
QY 66 VSPRNSPGEILLNLGELGDLNPLYLALSRMYNGYAGGVEQVLLAGNAFTAGKILFAAIP 125
Db 179 -STSTQKGILPKFQSLGPNLNLPHLEHLAKLYVWSSGSEVRSISGSGVFGKLAIVP 237
QY 126 PNFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFHFNNQOPQRMRLVAMLYTP 185
Db 238 PG--VDPVQSTMLQYPHVLPDARQVEPVIFCLPDLRSLTLXHLMSDTD--TTSLVIMVYND 294
QY 186 LRSNGSGDDVFTVSC--RVLTPTPDPEFIVLPP-----SVES-----KT----- 224
Db 295 L-INPYANDANSGCCIVTVETKPGPDFKHLKLPKGSMLTHGSIIPSDILPKTSSILWGNR 353
```

```
QY 225 -----KPTL-----PI-LTISELTNSRFPPIPIQLYTA 252
Db 354 YMSDITDVFIRFVQANHFDPNQETAGWSTPRFRPISVITEQNGAKLIGVATDIV 413
QY 253 PN-----ETNVVQCGNGRCTLDGBL-----QGTOLLSSAVCFLOGRTVAD---N 294
Db 414 PGIPDGPDT-----TIPGELIPAGDYAITNGTGNIDITAT---GYDTADIKN 459
QY 295 GDN-----W-DQNLQLLQTYPNGASVD--PTDEVPAPLGTQDFSGMLYGVLTQD 339
Db 460 NTNFRGMYICGSLQRAWGDKKISNTAFITATLDGNNKKNPCNTIDOSKI---VVFQD 516
QY 340 N-----VNVV-----TGEAKNAKG-----IYISTSGKFTPKIGSIGLHSTEHV 379
Db 517 NHVGKKAQTSDDTLALLGYTGIGEQAIGSDRDRVVRIST---LPETGARGN----- 565
QY 380 HPNQQRFTPVGVAVDENTPQQWVLPVHYAGSLALNTNLAP 420
Db 566 HPFYKNSIKLGVVIRSIDVFNSQLHTSRQLSLNHYLLP 605
```

RESULT 12

```
VCWMF9
coat protein - feline calicivirus (strain Japanese F4)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C:Accession: B40481
R:Tohya, Y.; Taniguchi, Y.; Takahashi, E.; Utagawa, E.; Takeda, N.; Miyamura, K.; Yamaza
Virology 183, 810-814, 1991
A:Title: Sequence analysis of the 3'-end of feline calicivirus genome.
A:Reference number: A40481; MUID:91306470; PMID:1853578
A:Accession: B40481
A:Molecule type: genomic RNA
A:Residues: 1-668 <TOH>
A:Cross-references: GB:D90357; NID:G221264; PIDN:BAA14371.1; PID:G221266
A:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,301,304,399,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted
```

```
Query Match 9.5%; Score 271.5; DB 1; Length 668;
Best Local Similarity 22.5%; Pred. No. 3.3e-12;
Matches 130; Conservative 77; Mismatches 193; Indels 179; Gaps 25;
```

```
QY 30 EPVAGASTAAPVAGQNTIIDPW-IRTNFVQAPNGEFTVSPRNSPGEILLNLGELGDLNPY 88
Db 145 EPSAQWSTAADMASGKSDSEWEAFFSFHTSVNW---STSETQKGILFKOSLGLPLNPY 200
QY 89 LAHLRMYNGYAGGVEQVLLAGNAFTAGKILFAAIPDNFLVDMISPAQITMLPHLIVDV 148
Db 201 LEHLRMYNGYAGGVEQVLLAGNAFTAGKILFAAIPDNFLVDMISPAQITMLPHLIVDV 258
QY 149 RTLEPIMTLPDVRNVFHFNNQOPQRMRLVAMLYTP-----RNSGSGDDVFTVSCR 202
Db 259 RQVEPVIFTPDLRSTLYHVMSDTD--TTSLVIMVYNDLINPYANDNSNGCIV-----TV 312
QY 203 LTRPTPDPEFIVLPP-----SVESK-----TKPFTL----- 229
Db 313 ETKPGPDFKHLKLPKGSVLTGHSIPSDILPKSSSLWNTGNRWITDITDFVIRPFVQANR 372
QY 230 -----PI-LTISELTNSRFPPIPIQLYTAFTNFTNVVQCGNGRCTLDGE 271
Db 373 HFDNQETAGWSTPRFRPITITISEKNGSKLIGVATDVIIPGIP----- 417
QY 272 LQGTOLLSSAVCFLOGRTVANGDNW--DQNLQLLQTYPNGASVDPTDEVPAPLGTQDFSG 330
Db 418 -----DGWPDTTIADKLIPAGD----- 434
QY 331 MLYGVLTQDNVNVSTGEAKNAKGIYISTSGKFTPKIGSI---GLHST----- 376
Db 435 --YSITTTGENDIKTAQADTAADVKNYTNFRGMYICGSLQRAWGDKKISNTAFITAIR 492
```


QY 377 --EHVHPNQSRFTPVGVADVENTPQQWV-----LPHYA--GSLAINTNLAPAV-- 422
Db 493 DGNIEKPSNTIDMTK--LAVYQDTHVQSVQTSDDTLALLGTYGTIGEAGSNDVRVRI 550
QY 423 -----APTTPGSQLFFRSRVCVQGLQODAFIDCLL---PQWVNHFYQEAAPSQADV 474
Db 551 SVLPAGARGGNHPIFYKNSIKLVIRSDVFNQSLHTSRLSLHY---LLPDS-F 606
QY 475 ALIRVYNPDGTGTLFEAKLHRSQ--FITVSHTG--AYPL 509
Db 607 AVYRII--DSNGSWFDIGIDSGFSGVGVSDIGKLEFPL 643

RESULT 13
A48562
coat protein - San Miguel sea lion virus (serotype 1)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: A48562
R:Neill, J.D.
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lions.
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Molecule type: genomic RNA
A:Residues: 1-702 <NEI>
A:CROSS-references: GB:M87481; NID:g334882; PIDN:AAAL6217.1; PID:g334884
A:Note: sequence extracted from NCBI backbone (NCBIN:113564, NCBI:P:113565)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.4%; Score 269.5; DB 1; Length 702;
Best Local Similarity 25.0%; Pred. No. 5e-12; Mismatches 169; Indels 97; Gaps 21;
Matches 108; Conservative 58

QY 8 AAPSDGAAALVPEGINMTPLPVPAGAS----IAAPVAGQTNIIID-PWIRTNFVQ---A 59
Db 151 AESDGPAGADIVTEQGTVQVQVPVPAQSAITLAAASTGKT--VDCEN--TTFSYHTA 206
QY 60 PNGETVSPRNSPGBIILNLELGPDLNLYLAHLSRMVYAGGVQVLLAGNAFTAGKI 119
Db 207 VNWSTT---EAQGKILFSLRSLPDLNLYLAHLSRMVYAGGVQVLLAGNAFTAGKI 119
QY 120 LFAAIPPNFLVDMISPAQITMLPHLIVDVTLEPIPTPLDVRNVFYHFNQPPRMLV 179
Db 263 AALIVPPG--IEPVESPTMLQVPHVLPDAROTEPVIFTIPDIRKTLXH-SMDDTDTTLV 319
QY 180 AMLYTPL-RSNGSGDDVFTVSCRVLTRTPDPFEITYLVPP-----SVESKTKP----- 226
Db 320 IMVYNELINPVEQSPKSCSITVETRSDSFTFSLAPGSLKHGSIPLSDIPRNSRH 379
QY 227 -----FTLPILTIS-----ELTNSRFPPIEQLYTAPNETNVVQNGRCT 267
Db 380 WGNRNMSTIDGFVVQPRVFNQSRNHRFDFTTGNSTP-----YIPIEVLEKLDRG--- 432
QY 268 LDGELQGTQLLSSAVCFLOGRVADNGNDWDLQITYPNGASYDPTDEVPAPLGTQD 327
Db 433 --GOYFKVTDTEKSLVPLGLP-----DGMPTTII-----PT-AMTASNGNYD 470
QY 328 FSGMLYGLVLTQDNNVNVSTGEAKNAKGIYISITTSKGTPKI--GSIGLHSITEHVHPNQSR 386
Db 471 YVVAEYRI-----TNGTHFGFYI---MGNLTKVKGSDNLGETQO-----TSRTL 514
QY 387 FTPVGVADVENT 398
Db 515 FASVGVNYKDQNT 526

RESULT 14
JQ2356

capsid protein - feline calicivirus (strain KCD)
C:Species: feline calicivirus
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: JQ2356
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable regions
A:Reference number: JQ2354; MUID:94065683; PMID:7504075
A:Accession: JQ2356
A:Molecule type: mRNA
A:Residues: 1-568 <SEA>
A:CROSS-references: GB:L09719; NID:g305107; PIDN:AAAL6487.1; PID:g305108
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein

Query Match 9.4%; Score 268.5; DB 2; Length 668;
Best Local Similarity 24.3%; Pred. No. 5.6e-12;
Matches 132; Conservative 81; Mismatches 222; Indels 109; Gaps 27;
QY 30 EPVAGASIAAPVAGQTNIIIDPW-IRTNFVQAPNGEFTVSPRNSPGEIILLNLELGPDLNPNY 88
Db 145 EPNAQMSTAADMATCKSVDSWEAFSPHTSVNW---STSETQKILFKQSLGLPNPY 200
QY 89 LAHLSRMVYAGGVQVLLAGNAFTAGKILFAAIPNPNFLVDMISPAQITMLPHLIVDV 148
Db 201 LEHLAKLYVAVWSGSIEVRSISGSGVFGKLAALIVVPEG--VDPVQSTSMQLQYPHVLFDA 258
QY 149 RTLPIPTPLDVRNVFYHFNQPPRMLVAMLYTLP-----RSNGSGDDVFTVSCRV 202
Db 259 RQVFPVFSIPDLRSTLYHLSMDTD-TTSLVMAYNDLINPYANDNSNGCIY-----TV 312
QY 203 LTRTPDPFEITYLVPP-----SVESKTKPFTLPI-----LTISELTNSRFPPIEQ 248
Db 313 ETKPGSDPRFHLKPPGSLVTHGSPVSDLI1PKTSSLWIGNRVYWTNITNPFVIRPPFVFOANR 372
QY 249 LYTAPNETNVVQNGR-CTLDGELQGTQL-LSSAVCFLOGRVADNGDNW-DQNLQL 305
Db 373 HFDFNQETAGWTPRFRPITITISEQGGTKLIGVATDYI---VPGPDGMPDPTTISEE 428
QY 306 TYPNG-----ASYDPTDVPAPLGTQDFSGM-LYGLVLTQDNNVNVSTGEAK 349
Db 429 LIPAGDVAYTNDIGNDITTPAGYDAADTIK---NNTNFRGMVYCGSLOR-----AWGDKK 480
QY 350 NAKGIYISTT---SGKFTP-----KIGSGLHSITEHVHPNQSR-----RFTPVG---V 392
Db 481 ISNTAFITATVKNKLPKSNMIDQTKIAVQDNHVGDKVQTSDDTLALLGYTGIGEQAI 540
QY 393 AVDENTPQQWVLPHYAGSLALNTNLAPAVAPTPFGEQLLFFRSRVCVQGLQODAFID 452
Db 541 GSDRDKVVRISVLPETGAR-----CGNHPIFYKNSIKLVYIRSDVFNPS 585
QY 453 CLL---PQWVNHFYQEAAPSQADVVALIRVYNPDGTRTLFEAKLHRSQ--FITVSHTG-- 505
Db 586 QILHTRSLQLSLHYL--LAPD--SFAVYRII--DSNGSWFDIGIDSGFSGVGVSDIGKL 639
QY 506 AYPL 509
Db 640 EFPL 643

RESULT 15
C48562
coat protein - San Miguel sea lion virus (serotype 4)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: C48562
R:Neill, J.D.
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lions.
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Accession: C48562

A:Molecule type: genomic RNA
A:Residues: 1-703 <NEI>
A:Cross-references: GB:M87482; NID:g334886; PIDN:AAAL6220.1; PID:g334888
A>Note: sequence extracted from NCBI backbone (NCBI:113567)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 8.4%; Score 240; DB 1; Length 703;
Best Local Similarity 29.3%; Pred. No. 88-10; Indels 16; Gaps 8;
Matches 63; Conservative 45; Mismatches 91; Mismatches 16; Gaps 8;

QY 8 AAPSNDGAASLVPEG----INETMPLEPVAGASIAAPVAGQTNIIDPWIRTNFVOAPNCE 63
DB 151 AESDGPGEAEIVTEEGTVQQQPAPAPTALATLATASTGK-SVEQEW-TFF--SYHTS 206
QY 64 FTVSPRNSPEILLNLLELQFDLNPYLHLRMYNGVAGVEVOVLLAGNAFTAGKILFAA 123
DB 207 INWSTVESQKILYQALNPSINPYLDHIAKLYSTWSGGIDVFTVSGSGVFGGKLAALL 266
QY 124 IPPNFIIVDMISPAQITMLPHLIVDVRTLPIVTPDVRNVEYHFNNOPQPMRLVAMLY 183
DB 267 VPEG--VEPIESVSMLOYPHLVFDARQTEPVFTIPDIRKTLFHSNDETD-TTKLVINPY 323
QY 184 TPLRSNGSGDDVFTVSCRVLTRPTDPDEFIYLVPP 218
DB 324 ----ENGV-ENKTTCSITVETREPSADFTFALLKPP 353

Search completed: June 1, 2004, 13:55:30
Job time : 12.4581 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:36:02 ; Search time 10.6124 Seconds
(without alignments)
4985.230 Million cell updates/sec

Title: US-09-926-799-10
Perfect score: 2906
Sequence: 1 MKMASNDRAPSNDGAANLVP.....GNQFYTLAPMGSGGQRRRAQ 550

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2020	69.5	542	S60616	capsid protein - h
2	1827.5	62.9	539	S40111	capsid protein - h
3	1141	39.3	530	B37471	capsid protein - N
4	1089.5	37.5	546	B37491	major capsid prote
5	319.5	11.0	576	A53982	capsid protein - E
6	293	10.1	702	A48562	coat protein - San
7	290	10.0	2344	RRWRH	genome polyprotein
8	288	9.9	2344	S55399	genome polyprotein
9	287.5	9.9	703	C48562	coat protein - San
10	284	9.8	2344	S64740	genome polyprotein
11	283	9.7	668	1 VCMWFC	coat protein - fel
12	280.5	9.7	671	1 VCMWFC	coat protein - fel
13	278.5	9.6	668	JQ2354	capsid protein - f
14	270.5	9.3	668	1 VCMWFF	coat protein - fel
15	262.5	9.0	668	JQ2356	capsid protein - f
16	129.5	4.5	2332	1 GNNYF	genome polyprotein
17	125.5	4.3	2194	1 GNNYF7	genome polyprotein
18	124.5	4.3	2333	1 GNNY2F	genome polyprotein
19	119.5	4.1	3085	2 T00327	polyprotein - infe
20	119	4.1	2647	2 A37098	gelation factor AB
21	118	4.1	2336	2 S37077	genome polyprotein
22	117.5	4.0	3375	2 T19821	hypothetical prote
23	116	4.0	733	2 JQ1892	capsid protein - f
24	115.5	4.0	490	2 T02411	cellulase (EC 3.2
25	115	4.0	733	2 JQ1891	capsid protein - f
26	111.5	3.8	2175	1 GNNYBE	genome polyprotein
27	111.5	3.8	2206	1 GNNY27	genome polyprotein
28	110.5	3.8	399	2 A11929	N-acetyl-glucosami
29	110.5	3.8	1204	2 S70393	pol polyprotein -

30	110.5	3.8	2332	1 GNNY4F	genome polyprotein
31	109.5	3.8	833	2 S45041	genome polyprotein
32	109.5	3.8	1199	1 GNMV1M	HIV-1 retropepsin
33	108.5	3.7	1204	2 S35475	pol polyprotein -
34	107	3.7	1011	1 GNNYC1	genome polyprotein
35	107	3.7	1658	2 D75489	hypothetical prote
36	106.5	3.7	614	2 S58306	WD-40 repeat regul
37	105	3.6	3566	1 A40701	tenascin-X precurs
38	104.5	3.6	470	2 D84863	hypothetical prote
39	104.5	3.6	805	2 JC7635	aryl hydrocarbon r
40	104	3.6	492	2 T01584	cellulase (EC 3.2
41	104	3.6	3191	2 T22945	hypothetical prote
42	103.5	3.6	3624	2 AD0835	large repetitive p
43	103	3.5	1197	2 T30581	neural cell adhesi
44	102	3.5	883	2 T51221	hypothetical prote
45	102	3.5	1184	2 H70761	probable dnapolyme

ALIGNMENTS

RESULT 1

S60616
capsid protein - human calicivirus (strain Melksham)
C:Species: human calicivirus
A:Variety: strain Melksham
C:Date: 23-May-1997 #sequence_revision 23-May-1997 #text_change 28-Jul-2000
C:Accession: S60616
R:Green, S.M.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
Virus Res. 37, 271-283, 1995
A:Title: Capsid diversity in small round-structured viruses: molecular characterization
A:Reference number: S60615; MUID:96136658; PMID:8533462
A:Accession: S60616
A:Molecule type: genomic RNA
A:Residues: 1-542 <GR>
A:Cross-references: EMBL:X81879; NID:G976077; PIDN:CAAS7462.1; PID:G976079
A:Experimental source: strain Melksham
A:Note: it is uncertain whether Met-1 or Met-3 is the initiator
C:Superfamily: human calicivirus capsid protein
C:Keywords: capsid protein; coat protein

Query Match 69.5%; Score 2020; DB 2; Length 542;
Best Local Similarity 69.3%; Pred. No. 5.1e-147;
Matches 383; Conservative 68; Mismatches 88; Indels 14; Gaps 5;

QY	1	MMKASNDRAPSNDGAANLVP	EADEVMAL	EPVVGASIAAPVVGQNIIDPWIRENFVQAP	60
DB	1	MMKASNDRAPSDGDAGLV	PSNEVMAL	EPVAGAAAPVTGQTNIIDPWIRANFVQAP	60
QY	61	QGFEVTPSRNSFGEMLLN	LELGP	ELNLYSLSRMYNGYAGGMVQVVLGNAFTAGKII	120
DB	61	NGEFTVSPRNAPGEVLLN	LELGP	ELNLYSLSRMYNGYAGGMVQVVLGNAFTAGKLV	120
QY	121	FAAVPPHPFVENISAAQIT	MCPTMCPH	IVDVRLPVLPLPDINRFRFHNQENTPMLVA	180
DB	121	FAAVPPHPFVENISLPKI	TNFPFHVID	VRTLEPVLPLPDVNRNSFFHNQKDDPMRIVA	180
QY	181	MLYTPLRAN-SGEDVFTV	SCRVLTRP	ADFFETFLVPPTVESKTKPFTLPILTGLSLS	239
DB	181	MLYTPLRNSGDDVFTV	SCRVLTRPS	PDFFTLYLPPTVESKTKPFTLPILTGLSLS	240
QY	240	RFPAAIDMLYTDPNESI	VVPQNGRC	TLDGTLGGTQTQLVPTQICAFRGTLISQTAARAS	299
DB	240	RFPAAIDMLYTDPNESI	VVPQNGRC	TLDGTLGGTQTQLVPTQICAFRGTLISQTAARAS	299
QY	299	RFPVDPIDQMTSPNEVI	SVCQNGRC	TLDGELGGTQTQLVSGICAFKGEV---TAHLHDN	297
DB	299	RFPVDPIDQMTSPNEVI	SVCQNGRC	TLDGELGGTQTQLVSGICAFKGEV---TAHLHDN	297
QY	300	TDSQORARNHPLHVQK	MDGTQVDP	TDIDPAVLGAIDFKGTVFQVAGSORDVSGQOEQH	359
DB	300	TDSQORARNHPLHVQK	MDGTQVDP	TDIDPAVLGAIDFKGTVFQVAGSORDVSGQOEQH	359
QY	359	-----DHLNVTIT	NLGSP	EDIPAPLVDPFQGRVFGVI	349
DB	359	-----DHLNVTIT	NLGSP	EDIPAPLVDPFQGRVFGVI	349
QY	360	YATRAHEAHIDTTPK	YAPKLGTL	IKS-GSDDFTNQPIRFTPPVGMGD-NNNRWQELPD	417
DB	360	YATRAHEAHIDTTPK	YAPKLGTL	IKS-GSDDFTNQPIRFTPPVGMGD-NNNRWQELPD	417
QY	417	PANRGHDAVPTTYTAQ	YTPKLG	IQIGTWQTDLLTVNQPVKFTPVGLNDTEHFNQWVPR	409
DB	417	PANRGHDAVPTTYTAQ	YTPKLG	IQIGTWQTDLLTVNQPVKFTPVGLNDTEHFNQWVPR	409

C;Accession: B37491
R;Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
Science 259, 516-519, 1993
A;Title: Sequence and genome organization of a human small round-structured (Norwalk-like virus) serotype
A;Reference number: A37491; MUID:93142023; PMID:8380940
A;Accession: B37491
A;Status: not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-546 <LAM>
A;Cross-references: GB:I04718; NID:g1236787; PIDN:AAA92984.1; PID:g295114
A;Note: sequence extracted from NCBI backbone (NCBIP:123458)
A;Note: small round-structured virus, SRV, Norwalk virus, Norwalk-like virus, serotype
C;Superfamily: human calicivirus capsid protein
C;Keywords: Glycoprotein
F;303,340,441/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 37.5%; Score 1089.5; DB 2; Length 546;
Best Local Similarity 42.6%; Pred No. 1.2e-75;
Matches 250; Conservative 75; Mismatches 173; Indels 89; Gaps 16;

Qy 1 MKMASNDAAPSNDGAA---NLVPEAND-EYMALEPVVGASIAAPVVGQQNIIDPWIRENF 56
| | | | | :
Db 1 MMWASKDAPQSADGASGAGQLVPVEWTADPLPMEPVAGPTTAVATAGOVNMIDPWIVNF 60
| | | | | :
Qy 57 VQAPQGSEFTVSPRNSPOEMLNLELGPELNPYLSHLSRMNGYAGGMQGVVLGNFTA 116
| | | | | :
Db 61 VQSPQGSEFTISPNTPGDIILFDLQGLHLPFLSHLSQMTYMGVGNRVILLAGNAFS 120
| | | | | :
Qy 117 GKIIFAAVPHEFPVENISAQITMCPHVIVDVQLEPVLPLDIRNRFHYNOENTPRM 176
| | | | | :
Db 121 GKIIIVCCVPGFTSSLTIAQAATLFPHVIADVRLTEIEMPLEEDVRNVLTHTN-DNQPTM 179
| | | | | :
Qy 177 RLVAMLYTPLRANGS---EDVFTSVSRVLTRPAPDFEFTFLVPPTVESKTKPFTLPILT 233
| | | | | :
Db 180 RLVCMLYTLRTGGSGSNSDSFVVAGRVLTA PSSDFSFLVLPVTIEQKTRAFTVNPI 239
| | | | | :
Qy 234 GEISNSRFPAAIDMYLTDPNESIVVQPONGRCTLDGTQTQLVTPQICAFRGTLISQT 293
| | | | | :
Db 240 QTLNSRRFSLIQMILLSPDASVQVQPQNGRCIIDQLGTTTATSQGLFRVRGK-INQG 298
| | | | | :
Qy 294 ARAADSTDSQRARNHPLHVQNKLDTQDYDDTDIPAVIGAIDF-----KG 340
| | | :
Db 299 ARTLNLTE-----VDGKPFMAFDS-PAPVGFPDFGKCDWHMRISKTPNN 341
| | | :
Qy 341 TVFG-----VASQRDVSQGQEQGHYA TRAHAETHDTPKYAPKGLTIKSGSDDFN-- 393
| | | :
Db 342 TGSQDPNRSVSVQTNVG-----FVPHLSIQF---DEVFNHP 376
| | | :
Qy 394 -----TNQIRTFPMGMDNNWRQNELPDYSGRLLTNLNNLA PAVSPSPFGERILF 443
| | | :
Db 377 TGDYIGTIEWISQP-STPPGTDIN---LWEIPDYSSLSQAANLPA VPVPPGFGEALVY 431
| | | :
Qy 444 FRSIVPSAGGYGS-GYIDCLIPQEWOHVFQEAPSO SAVALRYVYNPDGRIIFEA KHLH 502
| | | :
Db 432 FVSAFPQGNRRSAPNDVPCLLPQEIYTHFVSEQAPTMDGAALLHYVD PDNRLNGEPKLY 491
| | | :
Qy 503 REGFLTVA---NCGNNPTVWPNGVYFEAWGNOFYTLAPMGSGQGR 547
| | | :
Db 492 PGGYLTCTVPGVNGVGPPOOLPLNGVLFVSVWSYRQ LKPVGTA STAR 538
| | | :

RESULT 5
A53982
C;capsid protein - European brown hare syndrome virus
C;Species: European brown hare syndrome virus
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 28-Jul-2000
C;Accession: A53982
R;Wirblach, C.; Meyerz, G.; Ohlinger, V.F.; Capucci, L.; Eskens, U.; Haas, B.; Thiel, H.
J. Virol. 68, 5164-5173, 1994
A;Title: European brown hare syndrome virus: relationship to rabbit hemorrhagic disease
A;Reference number: A53982; MUID:94309183; PMID:7518531
A;Accession: A53982
A;Status: preliminary

Matches 117; Conservative 61; Mismatches 150; Indels 140; Gaps 21;

QY 8 AAPSNDGAANLVPEANDEVMALPEVVGAS-----IAAPVVGQNIIDPWIRENFVQAPQGE 63

Db 151 AESDGFQGADIIVTEBQGTWVQOQPVPAQSALTTLAASTG-----KTVDC 196

QY 64 FTV-----SPRNSPGEMLLNLELGPPELNPYLHLSRMYNGYAGGMQVQVVLGNA 113

Db 197 WTTFSYHTAVNWSITEAGKILFSRALSPELNPYLHLSRMYNGYAGGMQVQVVLGNA 256

QY 114 FTAGKIIFAAVPPHPPFVENISAAQITMCPHIVDVROLEPVLPLPDIRNRPFFHYNQ 173

Db 257 VEGKLAALIVPP--GIEVESFTMLQXPHVLFDARQTEPVFTIPDIPKTLIH-SMDDT 313

QY 174 PMRLVAMLYTPI--RANGSGEDVFTVSCRVLTRPAPDFEFTLVPTVESKTKPTLPI 231

Db 314 DTRLLVIMVYNELINDEYQSEPKSCSITVETRPSSDFTFLSKPP--GSLLKHSIPS 371

QY 232 TLGELS-----NSRFPAAIDMLYDTPNESIVQFQ-----NGRCTLGDTLQCTTQLVPTQIC 283

Db 372 LIPRNSRHWMGNRWSTID-----GFVQVQPRVQSNRHFDFDSTTTGWS--TFYI- 420

QY 284 AFRGTLISQTARAADSTDSQPARNHPLHVQVKNLD-GTQYD-----P 325

Db 421 -----PIEVTLKLDRGQYFKYVTTTEKSLVPLGPDGWP 454

QY 326 TDDIPAVL-----GAIDFKGTGVFVASQDVSQGOEQGHYATRAHEAHIIDTPKYPAPK 381

Db 455 DTPIPTAMTASNGYDVTAEYRITN-----NGTHFKGFYI-----MG 492

QY 382 TLIIK-SGSDPF-NTNQPIR--FTPVGMGNWQWQWELPDYSGLTLN 425

Db 493 NLATTKVKGSDNLGETQTSRTLSFVG-----NYKDQNTIN 528

RESULT 7

RRWRH

genome polyprotein - rabbit hemorrhagic disease virus

N;Contains: RNA-directed RNA polymerase (EC 2.7.7.48)

C;Species: rabbit hemorrhagic disease virus

C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 29-May-1998

C;Accession: A41039

R;Meyers, G.; Wirblich, C.; Thiel, H.J.

Virolgy 184, 664-676, 1991

A;Title: Rabbit hemorrhagic disease virus-molecular cloning and nucleotide sequencing

A;Reference number: A41039; MUID:91361557; PMID:1840711

A;Accession: A41039

A;Molecule type: genomic RNA

A;Residues: 1-2344 <MEY>

A;Cross-references: GB:M67473

C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase

C;Keywords: nucleotidyltransferase

Query Match 10.0%; Score 290; DB 1; Length 2344;

Best Local Similarity 30.3%; Pred. No. 2.1e-13;

Matches 89; Conservative 49; Mismatches 114; Indels 42; Gaps 12;

QY 8 AAPSNDGAANLVPEANDEVMALPEVVGAS-----IAAPVVGQNIIDPWIRENFVQAPQGE 63

Db 1772 AAPQGEAAGTATTASVPGTTDGMGPVVATTSTVTAENSASATAGIGGPPQVQDOQE 1831

QY 50 PHIRENFVQAPQGEFTVSPRNSPGEMLLNLELGPPELNPYLHLSRMYNGYAGGMQVQVVL 109

Db 1832 TW-RTNFY--NDVFTWSVADAPGSLITYVQHSPPQNPFATVLSQMYAGWAGMQFRFIV 1888

QY 110 AGNAFTAGKIIFAAVPPHPPFVENISAAQITMCPHIVDVROLEPVLPLPDIRNRPFFHYN 169

Db 1889 AGSGVFGRLVRAVIPP--GIEIGFGLVQRPFFHVVIDARSLEPVTITMPDLRPNMYHTP 1946

QY 170 QENTPRMLVAMLYTPIRANGSGEDVFTVSCRVLTRPAPDFEFTLVPP--TVESKTKP- 225

Db 1947 GDFGLVTLVLSVYNNLINPFGGSAIQVTVETRPSEDFEFVMIAPSSKTVDSISPA 2006

QY 226 -FTLPILTIGEISNSRFPAAIDMLYDTPNESIVQFQNG-----RCTLDGTLOCTT 275

Db 2007 LTTTPVLT-GVGNDRWNGQI-----VGLQPVPGGFTCNHRHWNLNGST 2049

RESULT 8

S55399

genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)

C;Species: rabbit hemorrhagic disease virus

A;Variety: isolate BS89

C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 18-Jun-1999

C;Accession: S55399

R;Roosi, C.

submitted to the EMBL Data Library, May 1995

A;Reference number: S55399

A;Accession: S55399

A;Molecule type: genomic RNA

A;Residues: 1-2344 <ROS>

A;Cross-references: EMBL:X87607; NID:g854640; PIDN:CAA60910.1; PID:g854641

A;Experimental source: isolate BS89

C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase

C;Keywords: polyprotein

Query Match 9.9%; Score 288; DB 2; Length 2344;

Best Local Similarity 30.4%; Pred. No. 3.1e-13;

Matches 89; Conservative 48; Mismatches 114; Indels 42; Gaps 12;

QY 9 APSNDGAANLVPEANDEVMALPEVVGAS-----IAAPVVGQNIIDPWIRENFVQAPQGE 63

Db 1773 AAPQGEAAGTATTASVPGTTDGMGPVVATTSTVTAENSASATAGIGGPPQVQDOQE 1832

QY 51 WIRENFVQAPQGEFTVSPRNSPGEMLLNLELGPPELNPYLHLSRMYNGYAGGMQVQVVL 110

Db 1833 W-RTNFY--NDVFTWSVADAPGSLITYVQHSPPQNPFATVLSQMYAGWAGMQFRFIV 1889

QY 111 GNAFTAGKIIFAAVPPHPPFVENISAAQITMCPHIVDVROLEPVLPLPDIRNRPFFHYN 170

Db 1890 GSGVFGRLVRAVIPP--GIEIGFGLVQRPFFHVVIDARSLEPVTITMPDLRPNMYHTP 1947

QY 171 QENTPRMLVAMLYTPIRANGSGEDVFTVSCRVLTRPAPDFEFTLVPP--TVESKTKP- 225

Db 1948 DGLVTLVLSVYNNLINPFGGSAIQVTVETRPSEDFEFVMIAPSSKTVDSISPA 2007

QY 226 FTLPILTIGEISNSRFPAAIDMLYDTPNESIVQFQNG-----RCTLDGTLOCTT 275

Db 2008 LTTTPVLT-GVGNDRWNGQI-----VGLQPVPGGFTCNHRHWNLNGST 2049

RESULT 9

C48562

coat protein - San Miguel sea lion virus (serotype 4)

N;Alternate names: capsid protein

C;Species: San Miguel sea lion virus

C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999

C;Accession: C48562

R;Neill, J.D.

Virus Res. 24, 211-222, 1992

A;Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lion virus

A;Reference number: A48562; MUID:92410750; PMID:1529644

A;Accession: C48562

A;Molecule type: genomic RNA

A;Residues: 1-703 <NEI>

A;Cross-references: GB:M87482; NID:g334886; PIDN:AAAL6220.1; PID:g334888

A;Note: sequence extracted from NCBI backbone (NCBI:P:113567)

C;Superfamily: feline calicivirus coat protein

C;Keywords: capsid protein; coat protein; glycoprotein

F;89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.9%; Score 287.5; DB 1; Length 703;

Best Local Similarity 24.5%; Pred. No. 5.9e-14;

Matches 104; Conservative 67; Mismatches 144; Indels 109; Gaps 18;

QY 8 AAPSNDGANLVPEANDEVMALPEVVG-----ASIAAPVVGQQNIIDPWIRENFVQAPQGE 63
Db 151 AESDGFSGSAEIVTEEQGTVVQQQPAPAPATLATATAGT-KSVEQEMWTFYSYHTSINW 209
QY 64 FTVSPRNSPGEMLNLLELGPENLPYLHSLRMVYAGGQVQVVLGNAFTAGKIIFAA 123
Db 210 STV---ESQKLYSQALNPSINPYLDHAKLYSTWSSGIDVRFTVSGSGVFGGKLAALL 266
QY 124 VPPHF--PVENISAAQITMCPHVIVDVRQLEPVLPLPDIRNFFHYNQENTPRMLVAML 182
Db 267 VPPGVEPIESVMLOQ---PHVLFDARQTEPIVFTIPDIRKTLFH-SMDETDTTKLVINP 322
QY 183 YTPLRANSGEDVFTVSCRVLTRPADPFEFTFLVPP----- 217
Db 323 Y-----ENGVENKTCISITVETRPSADFTFALLKPPGSLIKHGSIPSDLIPRNSAHMMGNR 378
QY 218 -----TVESKTKPFTLP--ILTLGEL-SNSRPPAAIDMLY 249
Db 379 WNSTISGFSVQPRVQSNRHFDFDSTTTCWSTPYVVPBIKQKRGVSNKRWHFVID--- 435
QY 250 TD-----PNESIVVQPNQGRCTLDGTLQ-GTTLVPTQICA-FRGTLLIS 291
Db 436 TDKALVPGIGDPDPTTIDET---KATNGNFSYSGSYRAGSTTIKPNENSTHKGTYIC 492
QY 292 QTARAADSTDSQPARANHPLHVQKMLDGTQDPTDDIPAVLGAIDFKGTVEFGVASQRDV 351
Db 493 GTLSTVEIPENDEQ---QIKTEAEKKSQTMVVT-----ADFKDTI--VKPQHKI 537
QY 352 SGQQ 355
Db 538 SPQK 541
RESULT 10
S64740
genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)
N:Contains: VP60 protein
C:Species: rabbit hemorrhagic disease virus
A:Variety: isolate AST/89
C:Date: 12-Jul-1996 #sequence revision 26-Jul-1996 #text_change 20-Jun-2000
R:Casais, R.; Martin-Alonso, J.; Boga, J.; Parra, F.
submitted to the EMBL Data Library, May 1995
A:Description: Genomic organization of rabbit hemorrhagic disease virus determined by di
A:Reference number: S64740
A:Accession: S64740
A:Molecule type: genomic RNA
A:Residues: 1-2344 <CAS>
A:Cross-references: EMBL:Z49271; NID:G1182032; PIDN:CAA89265.1; PID:G1150552
A:Experimental source: isolate AST/89
R:Boga, J.; Casais, R.; Martin-Alonso, J.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra, F.
submitted to the EMBL Data Library, July 1993
A:Description: Molecular cloning, sequence and expression of the capsid protein gene fro
A:Reference number: S46944
A:Accession: S46944
A:Molecule type: genomic RNA
A:Residues: 1650-2344 <BOG>
A:Cross-references: EMBL:Z24757; NID:G515622; PIDN:CAA80881.1; PID:G515623
A:Experimental source: isolate AST/89
R:Parra, F.; Boga, J.A.; Marin, M.S.; Casais, R.
Virus Res. 27, 219-228, 1993
A:Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus supp
A:Reference number: S49018; MUID:93255896; PMID:8488721
A:Accession: S49018
A:Molecule type: genomic RNA
A:Residues: 1650-1796 <PAR>
A:Cross-references: EMBL:Z24757
A:Experimental source: isolate AST/89
A:Accession: S65012
A:Molecule type: protein
A:Residues: 1767-1779;1875-1877,'X',1879-1881;1936-1938,'X',1940-1941 <PAW>
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: blocked amino end; polyprotein

Query Match 9.8%; Score 284; DB 2; Length 2344;
Best Local Similarity 30.4%; Pred. No. 6.2e-13;
Matches 89; Conservative 47; Mismatches 115; Indels 42; Gaps 12;
QY 9 AFSNDGANLV---VPEANDE-----VMALEPVV-----GASIAAPVVG-----QQNIIDP 50
Db 1773 APOGAAAGTATTASVPGTTTIDGMDFGVWATTSVWTAENSASATAGIAGPPQVDDQOET 1832
QY 51 WIRENFVQAPQGEFTVSPRNSPGEMLNLLELGPENLPYLHSLRMVYAGGQVQVVLGNAFTAGKIIFAA 110
Db 1833 W-RTNFFY--NDVFTWSVADAFGSILYTVQHSFQNNPFTAVLSQMYAGAGGMOFFIVA 1889
QY 111 GNAFTAGKIIFAAVPHPPPEVENISAAQITMCPHVIVDVRQLEPVLPLPDIRNFFHYNQ 170
Db 1890 GIGVFGRLVAIVIPP--GIEIGPGLVQRQPHVVIDARSLSPVTTITMDDLPRNMYHPTG 1947
QY 171 ENTPRMLVAMLYTPLRANSGEDVFTVSCRVLTRPADPFEFTFLVPP---TVESKTKP-- 225
Db 1948 DFGLVPTLLVSYNNLINPFGCGSTSAIQVTVETRESEDFEFYMIAPSSKTVDSISAGL 2007
QY 226 FTLPTLTGLGELSNRFPAAIDMLYTDPNESIVVQPNQ---RCTLDGTLQGT 275
Db 2008 LTPVLT-GVGNDRWNGOI-----VGLQVPVGGFSTCNRHNLNGST 2049
RESULT 11
VCWVFC
coat protein - feline calicivirus (strain Japanese F4)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C:Accession: B40481
R:Tohya, Y.; Taniguchi, Y.; Takahaeahi, E.; Utagawa, E.; Takeda, N.; Miyamura, K.; Yamaz
Virology 183, 810-814, 1991
A:Title: Sequence analysis of the 3'-end of feline calicivirus genome.
A:Reference number: A40481; MUID:91306470; PMID:1853578
A:Accession: B40481
A:Molecule type: genomic RNA
A:Residues: 1-668 <TOH>
A:Cross-references: GB:D90357; NID:G221264; PIDN:BAAL4371.1; PID:G221266
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,301,304,399,459,615/Binding site: carbohydrate (Aan) (covalent) #status predicted
Query Match 9.7%; Score 283; DB 1; Length 668;
Best Local Similarity 24.2%; Pred. No. 1.2e-13;
Matches 131; Conservative 78; Mismatches 198; Indels 134; Gaps 27;
QY 11 SNDGANLVPEANDEV--MALEPVVVGASIAAPVVGQQNIIDPWIRENFVQAPQGEFTVSP 68
Db 125 ADDGSIT-TPEQGTAVGVIAEPSAQMTAADMASGKSVDSM--EAFPSF-HTSVNWS 180
QY 69 RNSPGEMLNLLELGPENLPYLHSLRMVYAGGQVQVVLGNAFTAGKIIFAAVPPHF 128
Db 181 SETQKILFKQSLGPLNLPYLEHLSKLYVAMSGSIEVRFSISGSGVFGKLAALVVPVGV 240
QY 129 -PVENISAAQITMCPHVIVDVRQLEPVLPLPDIRNFFHYNQENTPRMLVAMLYTPL- 186
Db 241 DPVQSTSMLOQ---PHVLFDARQVEPIVFTIPDLRSTLYHV-MSDITDTTSLVIMYNDLI 296
QY 187 -----RANSGEDVFTVSCRVLTRPADPFEFTFLVPPTVESKTKPFTLPTLTGLGELSNR 241
Db 297 NPYANDSNSGCIIVTE---TKPGDPFKHLLKPPG-----SVLTHGSPSOLI 342
QY 242 PAADIM-----LYTDPNESIVVQPNQGRCTLDGTLQGTTLVPTQICAFRGTLLISQTARA 296
Db 343 PKSSSLWIGNRYWTDITD-FVIRP-----FVFAQNRH 373
QY 297 AD---STDSPQARHNLHVQKMLDGTQYDPTDDIPAVLGAIDFKGTVEFGVASQRDVSG 353
Db 374 FDFNQETAGWSTPRPRPRTITISEKNGSK-----LG-----IGVATDYIIPG 415

354 QOEQGHYATRAHEAHTDTPKYAPKL---GTLIKSG-SDDFNTNQPIRFTPVGMGNN 409
416 -----IPDGHFDIT---IADKLIPAGDYSITTEGNDIKTAQAYDTAAVVKNTTN 462
410 WRQWELPDYSGRLTNMNLAPAVSPFGRILFRSIVPSAGGYG-----SGYID--C 461
463 FR-----GMVICSQORAW-GDKKISNTAFITTAIRDGNKIPSNITDMTKL 508
462 LIPOEWQHFYQEAAPSQSAVALRVNPDTCRNIPEAKLHREGFLTV-----ANCNNP 516
509 AYQD--THVQEVQTSDDTLALLGYTG--IGEEAIGSNRDRVVRISVLPEAGARGGNHP 564
517 I 517
565 I 565

RESULT 12
VCWVF9
coat protein - feline calicivirus (strain F9)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C:Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text_change 28-Jul-2000
C:Accession: B43382; C45538; FQ0407; S23702
R:Cartier, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.; Turner, P.C.
Virology 190, 443-448, 1992
A:Title: The complete nucleotide sequence of a feline calicivirus.
A:Reference number: A43382; MUID:92410623; PMID:1529544
A:Accession: B43382
A:Molecule type: genomic RNA
A:Residues: 1-671 <CAR1>
A:CROSS-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879
R:Cartier, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.
Arch. Virol. 122, 223-235, 1992
A:Title: Identification and sequence determination of the capsid protein gene of feline
A:Reference number: A45538; MUID:92117861; PMID:1731695
A:Accession: C45538
A:Molecule type: genomic RNA; protein
A:Residues: 1-671 <CAR2>
A:CROSS-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879
A:Experimental source: strain F9
A:Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBI:77462)
R:Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.
J. Gen. Virol. 73, 2429-2433, 1992
A:Title: The cloning, sequencing and expression of a major antigenic region from the feline
A:Reference number: FQ0407; MUID:93019069; PMID:1402818
A:Accession: FQ0407
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 217-266, 'SI', 269-336, 'V', 338-395, 'S', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>
A:CROSS-references: PIDN:AA823553.1; PID:g257083
A:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.7%; Score 280.5; DB 1; Length 671;
Best Local Similarity 24.1%; Pred. No. 1.9e-13;
Matches 130; Conservative 74; Mismatches 207; Indels 129; Gaps 24;

11 SNDGAANLVPEANDEV--MALEPVVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSP 68
125 ADDGSIT-APEQGTVMVGVIAPSAQMSAAMATGKSVDSSEW--EAFPSF-HTSVNWSMT 180
69 RNSPGEMLNLELGPENLPYLSHLSRMVNGYAGGQVQVVLGNAFTAGKIIFAAVPPHF 128
181 SETQKILFKQSLGPLNLPYLSHLSRMVNGYAGGQVQVVLGNAFTAGKIIFAAVPPHF 240
129 -PVENISAAQITMCHPHVVDVQLEPVLPLPDINRFRFHYNQENTPRMRLVAMLYTPL- 186
241 DPVQSTSMLOQY---PHVLFAQKQVEPVFCLPDLSLYHL-MSDITDTSVIMVYNDLI 296
187 -----RANGSGEDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTPFTLITLIGELSNRFP 241

297 NPYANDANSSGCIIVTVE-----TKPDPDFKFLHKPPG-----SMLTHGSTPSDLI 342
242 PAA-----IDMLYTDPNESIVVQPNQGRCTLDGTLQTTQLVPTQICAFRGTLISQTARAA 297
343 PKTSSSLWGNRYWSDITDFVIRP-----FVQANRHF 374
298 D-----STDSPORARNHPLHVQKNLDCGTQYDPTDDIPAVLGAIDFKGTVFVASQORDVSQ 354
375 DFNQETAGWSTPRFRPISVITTEQNG-----AKLG-----IGVATDYIVPG- 415
355 QEQGHYATRAHEAHTDTPKYAPKLGTILIKSGS--DDFNTNQPIRFTPVGMGNNRWQ 413
416 -----IPDGPWPTTIPGELIPAGDYAITNGTNDITATGATYDADIKNNTFR-- 464
414 ELPDYSGRLLTNMNLAPAVSPSPGRILFRSIVPSA---GGYSGYIDC-LIPOEWV- 468
465 -----GMVICSQORAW-GDKKISNTAFITATLDGNNKINPCNTIDQSKIV 512
469 ----QHFYQEAAPSQSAVALRVY-----VNPDTGRNIFPEAKLHREGFLTVANCGNNPI 517
513 VFQDNHVGGKKAQTSDDTLALLGYTGIGEQAIGSDRDRVVRISTLPETG-----ARGGNHPI 568

RESULT 13
JQ2354
capsid protein - feline calicivirus (strain NADC)
C:Species: feline calicivirus
C:Date: 28-Aug-1985 #sequence revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: JQ2354
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable
A:Reference number: JQ2354; MUID:94065683; PMID:7504075
A:Accession: JQ2354
A:Molecule type: mRNA
A:Residues: 1-668 <SEA>
A:CROSS-references: GB:L09718; NID:g305104; PIDN:AAA16485.1; PID:g305105
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein

Query Match 9.6%; Score 278.5; DB 2; Length 668;
Best Local Similarity 24.1%; Pred. No. 2.7e-13;
Matches 125; Conservative 73; Mismatches 193; Indels 127; Gaps 22;

30 EPVVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSPRNSPGEMLNLELGPENLYL 89
145 EFSAQMSAAMATGKSVDSSEW--EAFPSF-HTSVNWSMTSETQKILFKQSLGPLNLYL 201
90 SHLSRMVNGYAGGQVQVVLGNAFTAGKIIFAAVPPHF-PVENISAAQITMCHPHVVDV 148
202 EHLKLYVAVSGSVSEVRSISGSGVFGKLAIVVPPGVDPVQSTSMLOQY---PHVLFA 258
149 ROLPEVLLPLPDINRFRFHYNQENTPRMRLVAMLYTPL-----RANGSGEDVFTVSCRVL 202
259 RQVDPVIFIPDLRSLYHL-MPDITDTSVIMVYNDLINPYANDNSNGCIVTVE---- 313
203 TRPADPFTFLVPPTVESKTPFTLITLIGELSNRFPAA-----IDMLYTDPNESIVV 258
314 TKPDPDFKFLHKPPG-----SMLTHGSVPSDLIPKSSSLWGNRYWSDITDFV 363
259 QPNQGRCTLDGTLQTTQLVPTQICAFRGTLISQTARAA-----STDSPORARNHPLHVQ 315
364 RP-----FVQANRHFNFQETAGMSAPRFRPITITI 395
316 KNLGDTQYDPTDDIPAVLGAIDFKGTVFVASQORDVSQEQGHYATRAHEAHTDTPDK 375
396 SESKGSK-----LG-----IGVATDYIVPG-----IPDGPWPTTIAE 427
376 YAPKLTILIKSGS--DDFNTNQPIRFTPVGMGNNRWQELPDYSGRLLTNMNLAPAVSP 434
428 DLTPAGDYAITSGNGNDITGSEYDSTEVIKNTNFR-----GMVICSQOR 474
435 SFGPERILFRSIVPSAGGYG-----SGYID---CLIPQEWQHFYQEAAPSQSAVALVR 486


```
Db 475 AW-GDKKISNTAFITAIKEGNKIRPSNTIDMTKIAYQD--THVGEVQTSDDALALIG 531
QY 487 Y-----VNPDGTGRNIFPEAKLHREGFLTVANGNNPI 517
Db 532 YTGIGEQAIGSDRDRVVRISVLPEVG---ARGGNHPI 565

RESULT 14
VCWVFF
coat protein - feline calicivirus (strain CFI/68 FIV)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Aug-1999
C:Accession: A40507; B40507; T09246
R:Neill, J.D.; Reardon, I.M.; Heinrichson, R.L.
J. Virol. 65, 5440-5447, 1991
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline calicivirus
A:Reference number: A40507; MUID:91374597; PMID:1716692
A:Accession: A40507
A:Molecule type: genomic RNA
A:Residues: 1-668 <NE1>
A:Cross-references: GB:M32819; NID:g323874; PIDN:AAA42925.1; PID:g323875
A:Accession: B40507
A:Molecule type: protein
A:Residues: 373-379; 403-419; 481-489; 560-566 <NE2>
R:Neill, J.D.
submitted to the EMBL Data Library, April 1998
A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.
A:Reference number: Z16626
A:Accession: T09246
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: genomic RNA
A:Residues: 1-668 <NE3>
A:Cross-references: EMBL:U13992; NID:g3056875; PIDN:AAC13993.1; PID:g537256
A:Experimental source: strain CFI/68 FIV
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.3%; Score 270.5; DB 1; Length 668;
Best Local Similarity 31.1%; Pred. No. 1.1e-12;
Matches 75; Conservative 41; Mismatches 94; Indels 31; Gaps 10;
QY 11 SNDGAANLVPEANDEV--MALEPVVVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSP 68
Db 125 ADDGSI-TPEQGTWVGVIAPENQAMSTADMATGKSVDSW--EAFPSF-HTSVNWS 180
QY 69 RNSPGEMLLNLELGPENLPYLPHLQLEPVLLPLPDINRPFHYNQENTPRMLVAMLYTPL- 186
Db 181 SETQKILFKQSLGPLNLPYLTHLAKLYVWSGSDVDRFSGVFGGKLAIVVPPGI 240
QY 129 -PVENISAAQITMCPHVIVDVRLQLEPVLLPLPDINRPFHYNQENTPRMLVAMLYTPL- 186
Db 241 DPVQSTSMQY---PHVLFDARQVPEVIFSPDLRSTLYHL-MSDITDTTSLVIMVYNDLI 296
QY 187 -----RANGSEDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPFTLPILTGLGELNSRF 241
Db 297 NPYANDSNSGCIIVTE----TKPGDFKFLHLLKPPG-----SNLTHGSIPLSDLI 342
QY 242 P 242
Db 343 P 343

RESULT 15
JQ2356
capsid protein - feline calicivirus (strain KCD)
C:Species: feline calicivirus
C>Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: JQ2356
R:Seal, B.S.; Kidpath, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
```

A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable regions
A:Reference number: JQ2354; MUID:94065683; PMID:7504075
A:Accession: JQ2356
A:Molecule type: mRNA
A:Residues: 1-668 <SEA>
A:Cross-references: GB:L09719; NID:g305107; PIDN:AAA16487.1; PID:g305108
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein

Query Match 9.0%; Score 262.5; DB 2; Length 668;
Best Local Similarity 30.7%; Pred. No. 4.6e-12;
Matches 74; Conservative 40; Mismatches 96; Indels 31; Gaps 10;
QY 11 SNDGAANLVPEANDEV--MALEPVVVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSP 68
Db 125 ADDGSI-TPEQGTWVGVIAPENQAMSTADMATGKSVDSW--EAFPSF-HTSVNWS 180
QY 69 RNSPGEMLLNLELGPENLPYLPHLQLEPVLLPLPDINRPFHYNQENTPRMLVAMLYTPL- 186
Db 181 SETQKILFKQSLGPLNLPYLTHLAKLYVWSGSDVDRFSGVFGGKLAIVVPPGI 240
QY 129 -PVENISAAQITMCPHVIVDVRLQLEPVLLPLPDINRPFHYNQENTPRMLVAMLYTPL- 186
Db 241 DPVQSTSMQY---PHVLFDARQVPEVIFSPDLRSTLYHL-MSDITDTTSLVIMVYNDLI 296
QY 187 -----RANGSEDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPFTLPILTGLGELNSRF 241
Db 297 NPYANDSNSGCIIVTE----TKPGDFKFLHLLKPPG-----SVLTHGSIPLSDLI 342
QY 242 P 242
Db 343 P 343

Search completed: June 1, 2004, 13:55:32
Job time : 12.6124 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 46.8188 Seconds
(without alignments)
3289.030 Million cell updates/sec

Title: US-09-926-799-1
Perfect score: 2896
Sequence: 1 MMASKDAPTMMGTSGAGQ.....YOLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2896	100.0	545	4	AAB49700 Small rou
2	1979	68.3	544	4	AAB49703 Small rou
3	1977.5	68.3	546	4	AAB49702 Small rou
4	1950	67.3	530	2	AAR50972 Norwalk v
5	1950	67.3	530	7	ADC72176 Norwalk v
6	1938	66.9	530	4	AAB49701 Small rou
7	1900	65.6	530	2	AAR57091 Small rou
8	1225	42.3	540	4	AAB49706 Small rou
9	1178.5	40.7	548	4	AAB49705 Small rou
10	1174.5	40.6	535	4	AAB49707 Small rou
11	1172.5	40.5	542	4	AAB49708 Small rou
12	1157.5	40.0	539	4	AAB49704 Small rou
13	1157	40.0	548	5	AAB49702 Norwalk v
14	1154.5	39.9	550	4	AAB49709 Small rou
15	1013	35.0	541	4	AAB49710 Small rou
16	321	11.1	579	2	AAM08143 RHDV caps
17	276	9.5	547	4	AAM50108 Feline ca
18	276	9.5	671	4	AAM50107 Feline ca
19	272.5	9.4	669	4	AAB67461 Amino aci
20	272	9.4	623	4	AAB47044 Feline ca
21	272	9.4	668	4	AAB67462 Amino aci
22	269	9.3	623	4	AAB47043 Feline ca
23	269	9.3	668	2	AAR10886 Feline ca
24	269	9.3	668	4	AAE04304 Feline ca
25	262.5	9.1	622	4	AAB47045 Feline ca

26	154	5.3	2206	2	AAR22210	Aar22210 True type
27	147	5.1	40	5	AAU91273	Aau91273 Norwalk v
28	136	4.7	40	5	AAU91274	Aau91274 Norwalk v
29	132.5	4.6	2209	1	AAP20037	Aap20037 Sequence
30	125.5	4.3	2179	1	AAP60243	Aap60243 Sequence
31	122.5	4.2	6310	6	ABU39869	Abu39869 Protein e
32	122	4.2	3263	6	ABU49739	Abu49739 Protein e
33	116.5	4.0	1707	4	ABG22165	Abg22165 Novel hum
34	114.5	4.0	1765	6	ABU20131	Abu20131 Protein e
35	113	3.9	1555	6	ABJ25640	Abj25640 Aspergill
36	113	3.9	1832	6	ABJ26240	Abj26240 Aspergill
37	112.5	3.9	2164	1	AP81045	Ap81045 Sequence
38	112.5	3.9	2164	1	AP80131	Ap80131 Peptides
39	112.5	3.9	2914	5	ABP70089	Abp70089 Human NOV
40	112	3.9	423	7	ABR82681	AbR82681 Human mit
41	112	3.9	3930	6	ABU18893	Abu18893 Protein e
42	110.5	3.8	5291	7	ADC01014	Adc01014 Enterohae
43	110	3.8	1037	4	ABH71260	Abh71260 Drosophil
44	110	3.8	1234	4	ABH68510	Abh68510 Drosophil
45	110	3.8	2016	4	ABB63911	Abb63911 Drosophil

ALIGNMENTS

RESULT 1
AAB49700
ID AAB49700 standard; protein; 545 AA.
XX
AC AAB49700;
XX
DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 1.
XX
KW Small round structured virus; SRSV; food poisoning.
XX
OS Small round structured virus.
XX
PN WO200079280-A1.
XX
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000WO-JP004095.
XX
PR 22-JUN-1999; 99JP-00175928.
XX
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.

Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
WPI, 2001-080848/09.
N-FSDB; AAF29141.
Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.
Claim 1; Page 40-42; 84pp; Japanese.
This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks
Sequence 545 AA;
Query Match 100.0%; Score 2896; DB 4; Length 545;
Best Local Similarity 100.0%; Pred. No. 9,7e-250;

Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.

claim 1: page 47-49: 84pp: Japanese.

This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAR49700 - AAR49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

Sequence 544 AA;

Query Match 68.3%; Score 1979; DB 4; Length 544;
Best Local Similarity 66.8%; Pred. No. 9.3e-168;
Matches 368; Conservative 68; Mismatches 101; Indels 14; Gaps 6;

```

QY 1 MMWASKADPTNMDGTSAGOLVPEANTABPIESMEPVAGATAAAATAGQVNMIDPJMNNY 60
Db 1 MMWASKADPTSDAGTATGAGQLVPEVNTADPIIDPVAGSSTALATAGQVNLIDPWIINN 60

QY 61 VQAPQGEFTISPNNTGDIIFDLQLGPHLNPFLSHLAQMYNGWGMVKVLLAGNAFTA 120
Db 61 VQAPQGEFTISPNNTGCDVLFDLQLGPHLNPFLSHLSQMYNGWGMVRVVLVLAGNAFTA 120

QY 121 GKIIISCIIPGFAAQNISIAQATMFPHVIADVRVLEPIEVLPELDVRNVLPHNNNDAPTWR 180
Db 121 GKVIICVPPGFSQRTLSIAQATLFFPHVIADVRTLDPEVLELDVRNVLPHNNDTQPTWR 180

QY 181 LVCMLYTPLRASGSSSTGDPFVIAGRVLTCPSPDFSLFLVPNNVEQTKPFSPVNLPLN 240
Db 181 LLCMLYTPLRTGASGSGTDSFVVAGRVLTCPGPDFNFELFLVPTVEQKTRPFTVPNIPLK 240

QY 241 TLSNSRVPSLIIKGMVSRDRHQVQFQNGRVTLDDQLQGTTPTSASOLCKIRGSVFHANG 300
Db 241 YLSNSRIPNIEGNSLSPDQNTQNVQFQNGRCTIDGQPLGTTPTVSVSOLCKFRGRI--TSG 298

QY 301 GNGYNLTDLGSPYHAFESAPTFGPDLDGSCDWHMEAS--PTTQNTGDIVIKQINVKQES 358
Db 299 QRVLNLTDLGSPPMFAAPAPAGFPDLGSCDWHIEMSKIPNSSTQNNPIVTVNSVKPNSQ 358

QY 359 AFAPHLGTIQADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPVWVIPRYGTSITTEAAQLA 418
Db 359 QFVPHLLSSITLD--ENVSSSGDYIGTTQWITSPPSDSGGANTNFWKIPDYGSSLAEASQLA 416

QY 419 PPIYPGFGGEAIVPFMSDFPIAHGTNGLS-----VPTCIPOFVTHFVNEQAPRTGEAALL 474
Db 417 PAVYPPGFGNEVIVYFMSAIP---GPNQSGSPNLVPCLLPQEYITHFISEQAPIQGEAALL 473

QY 475 HVLDPDTHRNIGEFKLYPEGFMTCVPNMSGTGTPQTLPLNGYVVFVSWVSRFYOLKPVGTA 534
Db 474 HVLDPDTHRNIGEFKLYPGGYLTCVPNMSGTGPOQLPLDGVFVFASWVSRYFYQLKPVGTA 533

QY 535 GBA-CRLGIRR 544
Db 534 GPARGLGVR 544

```

RESULT 3

AAB49702

AAB49702:

04-APP-2001 (first entry)

2025 RELEASE UNDER E.O. 14176

food poisoning.

XX KW

OS Small round structured virus.
XX WO200079280-A1.
XX PD 28-DEC-2000.
XX PF 22-JUN-2000; 2000WO-JP004095.
XX PR 22-JUN-1999; 99JP-00175928.
XX PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX PA (DENK-) DENKA SEIKEN KK.
XX PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX DR WPI; 2001-080848/09.
XX DR N-PSDB; AAF29143.
XX PT Kit for the detection and typing of small round-structured virus (SRSV)
XX PT strains for investigation of food poisoning outbreaks, contains
XX PT antibodies.
XX PS Claim 1; Page 45-47; 84pp; Japanese.
XX CC This invention relates to a kit for the detection and typing of small
XX CC round structured virus (SRSV) strains. The kit contains antibodies
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX CC used for detecting and typing strains of SRSV in order to prevent the
XX CC spread of infection and to examine the epidemiology of outbreaks
XX SQ Sequence 546 AA;
Query Match 68.3%; Score 1977.5; DB 4; Length 546;
Best Local Similarity 66.6%; Pred. No. 1.3e-167;
Matches 367; Conservative 76; Mismatches 95; Indels 13; Gaps 7;
Qy 1 MMASKADPTNMDTSGAGQLVPEANTAEPTSMPEVAGAAATAAGOVNMDPWIMNMY 60
Db 1 MMASKADQASDAGSAGQLVPEVNTADPLMEPVAGPTTAVATAGOVNMDPWVNNF 60
Qy 61 VQAPQGETTISPNNTPGDILFDLQGLPHNPFSLHQAQMYNGWGNMKVLLAGNAFTA 120
Db 61 VQSPQGETTISPNNTPGDILFDLQGLPHNPFSLHQAQMYNGWGNMKVLLAGNAFTA 120
Qy 121 GKIIISCIIPGFAAQNISIAQATMFPVIAADVRLIEVPLEVDVRLVFNHNDNAPTMR 180
Db 121 GKIIIVCCVPPGTTSSSLTIAQATLFPVIAADVRLIEVPLEVDVRLVFNHNDNAPTMR 180
Qy 181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFVPPNVEQTKPFSPVNLPLN 240
Db 181 LVCMLYTPLRTGGSGNSDSFVAGRVLTAPSSDFSLFVPPNVEQTKRAFTVNLPLQ 240
Qy 241 TLSNRSVPSLIKSMVSRDHQGMQVQNGRVTLQGLQGTTPTSASQCKIRGVSFVHANG 300
Db 241 TLSNRSVPSLIQMLSPDASQVQNGRVTLQGLQGTTPTSASQCKIRGVSFVHANG 300
Qy 301 GNGVNLTLDCSPHYAFESPAPIGPDDLGECDWHEASPT-TQNTGTVIQLINVKQE-S 358
Db 299 ARTNLTLTEVDGKPFMAFSDPAPVGPDPFGKCDWHRISKTNNNTSSGDMRSVSQTNVQ 358
Qy 359 APAPHLGTIQADGLSDVSNNTMLAKLWVSPVSDGHRGDVDPVPIPRYGSTLLEAAQLA 418
Db 359 GFVPHLSIQDEVPNHTG-DYIGTIEWISQPTPGCTDNLNHEIPDYGSSLSQAANLA 417
Qy 419 PPIVPPGGEAIVFPMSPDPIAHNGNLS-VPCTIPQBFVTHFVNEQAPTRGEAALL 474
Db 418 PPIVPPGGEALVYFVSAPF--GNPNSAPNDVPCLLPQEVYTHFVSEQAPTMGDAALL 474
Qy 475 HYLDPDTHRNLFGEFLYEGEFTWCPVNSSGTPQTLPIGCVFVSVWSRYQLKPVGTA 534
Db 475 HYVDPDTHRNLFGEFLYEGEFTWCPVNSSGTPQTLPIGCVFVSVWSRYQLKPVGTA 534

Qy 535 GPA-CRLGIRR 544
Db 535 STARSRLGVRR 545
RESULT 4
AAR50972
ID AAR50972 standard; protein; 530 AA.
XX AC AAR50972;
XX DT 16-OCT-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 05-OCT-1994 (first entry)
XX DE Norwalk virus strain 8FIIa protein (encoded by ORF2).
XX KW Norwalk virus; pathogen; acute gastroenteritis; food poisoning;
XX KW seafood contamination; diagnostic assay; calcivirus; small round virus.
XX OS Norwalk virus; (strain 8FIIa).
XX PN WO9405700-A2.
XX PD 17-MAR-1994.
XX PF 07-SEP-1993; 93WO-US008447.
XX PR 07-SEP-1992; 92US-00941365.
XX PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX PI ~~MAISON DOG - ESTES MK - Ulang X Graham-DY~~
XX DR WPI; 1994-101125/12.
XX DR N-PSDB; AAQ56826.
DNA from Norwalk and related viruses - used for preparing prods. for use
in diagnostic assays, detection and vaccines for Norwalk and related
viruses.
XX Claim 14; Page 68-70; 156pp; English.
XX The Norwalk virus was isolated from stool samples from adult volunteers
infected with safety tested Norwalk virus strain 8FIIa. The coding
sequence is useful for the design of probes for use in diagnostic assays
for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct
FN field.) (Updated on 16-OCT-2003 to standardise OS field)
XX SQ Sequence 530 AA;
Query Match 67.3%; Score 1950; DB 2; Length 530;
Best Local Similarity 67.0%; Pred. No. 3.5e-165;
Matches 366; Conservative 66; Mismatches 96; Indels 18; Gaps 7;
Qy 1 MMASKADPTNMDTSGAGQLVPEANTAEPTSMPEVAGAAATAAGOVNMDPWIMNMY 60
Db 1 MMASKADTSSVDASGAGQLVPEVNASDPLAMPVAGSTAVATAGOVNMDPWVNNF 60
Qy 61 VQAPQGETTISPNNTPGDILFDLQGLPHNPFSLHQAQMYNGWGNMKVLLAGNAFTA 120
Db 61 VQAPQGETTISPNNTPGDILFDLQGLPHNPFSLHQAQMYNGWGNMKVLLAGNAFTA 120
Qy 121 GKIIISCIIPGFAAQNISIAQATMFPVIAADVRLIEVPLEVDVRLVFNHNDNAPTMR 179
Db 121 GKIIIVCCVPPGTTSSSLTIAQATLFPVIAADVRLIEVPLEVDVRLVFNHNDNAPTMR 180
Qy 180 RLVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFVPPNVEQTKPFSPVNLPL 239
Db 181 RLVCMLYTPLRTGGTG--DSFVAGRVMTCPSPDFSLFVPPNVEQTKPFSLP 238
Qy 240 NTLNSRVPSLIKSMVSRDHQGMQVQNGRVTLQGLQGTTPTSASQCKIRGVSFVHANG 299

Db	239	SSLSNSRAPLPISSMGISPDNVQVQFQNGRCTLGRLVGTTPVSLSHVAKIRGT-----S	294
Qy	300	GGNGYNLTDLGSPYHAFESPAPIGFDPDLGECDHMEASPTQNTGDKVTKQINVKQESA	359
Db	295	NGTVINLTDLGTPPHPEGPAPIGFDPDLGECDHIN-----MTQFGHSSQTYQYDVTTPDT	351
Qy	360	FAPHLGTIQADGLSDVSVNTNMIKLGWVSPVSDGHRGDVDPWVIPIRYGSTLTTEAAQLAP	419
Db	352	FVPHLGSIQANGIG-----SGNYVGVLSWISPPSPSGSQVDLWKIPNYGSSITEATHLAP	407
Qy	420	PIYPGFGGALVFFMSDPPIAHGTNGLSVPCCTIIOEFTHFWNQAPTRGEEAALLHYLDP	479
Db	408	SVYPPGFGELVFFMSKMP-----GPGAYNLPCLLPQEIYISHLASQAPTVGGEAALLHYVDP	464
Qy	480	DTHRNLGFEKLYPEGFMTCVNSSGTGPQLPENGVFVSVSRFYOLKPVGTAGPA-C	538
Db	465	DTGRNLGFEKAYPDGFLTCVENGASSGQQLPINGVFVSVSRFYOLKPVGTASSARG	524
Qy	539	RLGIRR 544	
Db	525	RLGLRR 530	
RESULT 5			
ID	ADC72176	standard; protein; 530 AA.	
XX	AC	ADC72176;	
XX	DT	18-DEC-2003 (first entry)	
XX	DE	Norwalk virus protein 2 amino acid sequence.	
XX	KW	immune response; non-Norwalk virus agent; immunogen; Norwalk virus;	
XX	KW	viral pathogen; acute gastroenteritis; virucidal; antiviral vaccine.	
XX	OS	Norwalk virus.	
XX	PN	US6572862-B1.	
XX	PD	03-JUN-2003.	
XX	PF	07-JUN-1995; 95US-00486049.	
XX	PR	08-NOV-1989; 89US-00433492.	
XX	PR	27-APR-1990; 90US-00515993.	
XX	PR	27-AUG-1990; 90US-00573509.	
XX	PR	06-MAY-1991; 91US-00696454.	
XX	PA	(BAYU) BAYLOR COLLEGE MEDICINE.	
XX	PI	ates MK, Jiang X, Graham DY; ;	
XX	DR	WPI; 2003-776005/73.	
XX	DR	N-PSDB; ADC72174.	
XX	PT	Inducing an immune response against non-Norwalk virus agents, comprises	
XX	PT	administering an immunogen recombinantly expressed from a cDNA from	
XX	PT	Norwalk virus.	
XX	PS	Example 4; SEQ ID NO 3; 45pp; English.	
XX	CC	This invention relates to a novel method of inducing an immune response	
XX	CC	in an individual against Norwalk virus and non-Norwalk virus agents, by	
XX	CC	orally or parenterally administering an immunogen recombinantly expressed	
XX	CC	or synthesised from a cDNA of Norwalk virus given in the specification.	
XX	CC	Norwalk virus is one of the most important viral pathogens, causing acute	
XX	CC	gastroenteritis. The invention may be used for the development of	
XX	CC	compounds with virucidal activity or an antiviral vaccine. The present	
XX	CC	sequence is the amino acid sequence of a protein encoded by the Norwalk	
XX	CC	virus genome of the invention.	
XX	PI	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;	
SQ Sequence 530 AA;			
Query Match 67.3%; Score 1950; DB 7; Length 530;			
Best Local Similarity 67.0%; Pred. No. 3.5e-165; Indels 18; Gaps 7;			
Matches 366; Conservative 66; Mismatches 96;			
Qy	1	MMASKADAPTNMDGTSGAGQLVPEANTAEPIISMPEFVAGATAAATAGQVNMIDPMNMY	60
Db	1	MMASKADATSSVDGASGAGQLVPEYNASDPLAMDPAVAGSTAVATAGQVNPIDPMIINF	60
Qy	61	VQAPQGEFTISPNTPGDIILFDLQGLPHLNPFLSHLAQMYNGWGVGMKVVLGNAFTA	120
Db	61	VQAPQGEFTISPNTPGDIILFDLQGLPHLNPFLSHLAQMYNGWGVGMKVVLGNAFTA	120
Qy	121	GKIIISCIIPGFAAQNIATAQTMPPHVTADVRLEPIEVPLEDVRNVLFNHND-NAPTM	179
Db	121	GKIIISCIIPGFGSHMLTIAQATLPHVJADVRLDPIEVPLEDVRNVLFNHNDNQOTM	180
Qy	180	RLVCMLYTTLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNNVEQTKFSPVNLPL	239
Db	181	RLVCMLYTTLRTGGTG--DSFVWAGRVMTCPSPDFNLFVLPPTVEQKTRPFTLNLPL	238
Qy	240	NTLSNRVPSLIKSMVSRDHQMVQFONGRVTLGQLOGTTPTASQLCKIRGVSVEHAN	299
Db	239	SSLSNSRAPLPISSMGISPDNVQVQFQNGRCTLGRLVGTTPVSLSHVAKIRGT-----S	294
Qy	300	GGNGYNLTDLGSPYHAFESPAPIGFDPDLGECDHMEASPTQNTGDKVTKQINVKQESA	359
Db	295	NGTVINLTDLGTPPHPEGPAPIGFDPDLGECDHIN-----MTQFGHSSQTYQYDVTTPDT	351
Qy	360	FAPHLGTIQADGLSDVSVNTNMIKLGWVSPVSDGHRGDVDPWVIPIRYGSTLTTEAAQLAP	419
Db	352	FVPHLGSIQANGIG-----SGNYVGVLSWISPPSPSGSQVDLWKIPNYGSSITEATHLAP	407
Qy	420	PIYPGFGGALVFFMSDPPIAHGTNGLSVPCCTIIOEFTHFWNQAPTRGEEAALLHYLDP	479
Db	408	SVYPPGFGELVFFMSKMP-----GPGAYNLPCLLPQEIYISHLASQAPTVGGEAALLHYVDP	464
Qy	480	DTHRNLGFEKLYPEGFMTCVNSSGTGPQLPENGVFVSVSRFYOLKPVGTAGPA-C	538
Db	465	DTGRNLGFEKAYPDGFLTCVENGASSGQQLPINGVFVSVSRFYOLKPVGTASSARG	524
Qy	539	RLGIRR 544	
Db	525	RLGLRR 530	
RESULT 6			
ID	AAB49701	standard; protein; 530 AA.	
XX	AC	AAB49701;	
XX	DT	04-APR-2001 (first entry)	
XX	DE	Small round structured virus protein (SEQ ID 2).	
XX	KW	Small round structured virus; SRSV; food poisoning.	
XX	OS	Small round structured virus.	
XX	PN	WO200079280-A1.	
XX	PD	28-DEC-2000.	
XX	PF	22-JUN-2000; 2000WO-JP004095.	
XX	PR	22-JUN-1999; 99JP-00175928.	
XX	PA	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.	
XX	PA	(DENK-) DENKA SEIKEN KK.	
XX	PI	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;	

XX WPI; 2001-080848/09.
 DR N-PSDB; AAF29142.
 XX
 XX
 PT Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX
 XX
 PS Claim 1; Page 42-45; 84pp; Japanese.
 XX
 XX This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX
 XX Sequence 530 AA;

Query Match 66.9%; Score 1938; DB 4; Length 530;
 Best Local Similarity 67.0%; Pred. No. 4.1e-164;
 Matches 366; Conservative 63; Mismatches 99; Indels 18; Gaps 7;
 Qy 1 MMASKDAPTNDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGOVNMDPWIMNYY 60
 Db 1 MMASKDAPTNDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGOVNMDPWIMNYY 60
 Qy 61 VOAPQGETTSPNNTPGDILFDLQGLPHLNFLLSHLAQYNGWGNMKVKVLLAGNAFTA 120
 Db 61 VOAPQGETTSPNNTPGDILFDLQGLPHLNFLLSHLAQYNGWGNMKVKVLLAGNAFTA 120
 Qy 121 GKIIISICPPGFAAQNISIAQATMPHVIADRVLEPIEVLVDVRLVFNHND-NAPT 179
 Db 121 GKIIISICPPGFAAQNISIAQATMPHVIADRVLEPIEVLVDVRLVFNHND-NAPT 179
 Qy 180 RLVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVKQTKPFVSNLPL 239
 Db 181 RLVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVKQTKPFVSNLPL 238
 Qy 240 NTLNSRVPSLISKMMVSRDHGQVQFONGRVTLTGQGLQGTTPTSASOLCKIRGSVFHAN 299
 Db 239 SSLNSRAPLPISGMGISPDNVQSVQFONGRVTLTGQGLQGTTPTSASOLCKIRGSVFHAN 294
 Qy 300 GGNGYNTLTDGSPYHAFESPAPIGFDPDLGCDWHMEASPTTQFNTGDKVQINVKQESA 359
 Db 295 NGTVINLTDLGTFPHFEPGAPIGFDPDLGCDWHMEASPTTQFNTGDKVQINVKQESA 351
 Qy 360 FAPHLGTIQADGLSDSVSNTHMIKLGWSPVSDGHRGVDVDPWVPRYGSTLTAAQIAP 419
 Db 352 FVPHLGSIQANGIG---SGNYIGVLSWVSPSPHPSGQVLDLWKIPNYGSSITEATHLAP 407
 Qy 420 PIYPRGGEALVPMSPDPIAHGNGLSVPCITIQEFVTHFVNEQAPTRGEAALHLYDP 479
 Db 408 SVYPPGGEALVPMSPDPIAHGNGLSVPCITIQEFVTHFVNEQAPTRGEAALHLYDP 464
 Qy 480 DTHNLGEFLYPPGFMTCVPSNNGTGPOTLPINGVFVSWSRFVQKLVGTAGPA-C 538
 Db 465 DTGRTLEFKAYPGFLTCVENGASSGQQLPINGVFVSWSRFVQKLVGTAGPA-C 524
 Qy 539 RLGIIR 544
 Db 525 RLGIIR 530

RESULT 7
 ID AAR57091
 XX AAR57091 standard; protein; 530 AA.
 AC AAR57091;

XX 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)

05-OCT-1994 (first entry)
 Small round virus SRSV/KY/89 capsid protein.
 pathogen; acute gastroenteritis; food poisoning; seafood contamination;
 diagnostic assay; human calicivirus; small round virus; SRSV; KY89;
 Norwalk virus; capsid protein.
 Small round structured virus.
 WO9405700-A2.
 17-MAR-1994.
 07-SEP-1993; 93WO-US008447.
 07-SEP-1992; 92US-00941365.
 (BAYU) BAYLOR COLLEGE MEDICINE.
 Maseon DO, Eileen MK, Jifang X, Graham DY,
 WPI; 1994-101125/12.
 N-PSDB; AAO56832.
 DNA from Norwalk and related viruses - used for preparing prods. for use
 in diagnostic assays, detection and vaccines for Norwalk and related
 viruses.
 Example 7; Fig 13a; 156pp; English.
 The known sequence for Norwalk virus was used to obtain the sequence of
 other Norwalk-related viruses such as SRSV/KY/89, an agent from a stool
 from an outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide
 cDNA sequence includes part of the polymerase region and the capsid
 region of the genome; the deduced amino acid sequences are AAR57092 and
 AAR57091, respectively. Expression of fragments and derivs. of Norwalk-
 related viruses permits development of diagnostic assays to detect
 antibodies, antigens, viral genetic material or antivirals. (Updated on
 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS
 field.)
 Query Match 65.6%; Score 1900; DB 2; Length 530;
 Best Local Similarity 65.7%; Pred. No. 1e-160;
 Matches 362; Conservative 66; Mismatches 95; Indels 28; Gaps 9;
 Qy 1 MMASKDAPTNDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGOVNMDPWIMNYY 60
 Db 1 MMASKDAPTNDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGOVNMDPWIMNYY 60
 Qy 61 VOAPQGETTSPNNTPGDILFDLQGLPHLNFLLSHLAQYNGWGNMKVKVLLAGNAFTA 120
 Db 61 VOAPQGETTSPNNTPGDILFDLQGLPHLNFLLSHLAQYNGWGNMKVKVLLAGNAFTA 120
 Qy 121 GKIIISICPPGFAAQNISIAQATMPHVIADRVLEPIEVLVDVRLVFNHND-NAPT 179
 Db 121 GKIIISICPPGFAAQNISIAQATMPHVIADRVLEPIEVLVDVRLVFNHND-NAPT 179
 Qy 180 RLVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVKQTKPFVSNLPL 239
 Db 181 RLVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVKQTKPFVSNLPL 238
 Qy 240 NTLNSRVPSLISKMMVSRDHGQVQFONGRVTLTGQGLQGTTPTSASOLCKIRGSVFHAN 299
 Db 239 SSLNSRAPLPISGMGISPDNVQSVQFONGRVTLTGQGLQGTTPTSASOLCKIRGSVFHAN 294
 Qy 300 GGNGYNTLTDGSPYHAFESPAPIGFDPDLGCDWHMEASPTTQFNTGDKVQINVKQIN 354
 Db 295 NGTVINLTDLGTFPHFEPGAPIGFDPDLGCDWHMEASPTTQFNTGDKVQINVKQIN 353
 Qy 355 QKESAFAPHLGTIQADGLSDSVSNTHMIKLGWSPVSDGHRGVDVDPWVPRYGSTLTLEA 414

Db 354 -----PHLGSIQANGIG-----SGNIGVLSWSPSPSHSGSQVDLWKIPNYGSSITEA 402
Qy 415 AQLAPPYPPGGEAIVFFMSDFPIAHGTNGLSVPTIPQEFVTHFVNEQAPTRGEAALL 474
Db 403 THLAPSVYSPGGEVLVFFMSKIP---GPGGDSLPCILLPQGYIHLASEQAPTVGEGPLL 459
Qy 475 HYLDPDTHRNIGBKLYPEGEMTCVPNSSGTPGOTPLPINGVFFVSVWSRFFYQLKPVGTA 534
Db 460 HYVDPDTHRNIGBKLYPEGEMTCVPNSSGTPGOTPLPINGVFFVSVWSRFFYQLKPVGTA 519
Qy 535 GPA-CRLGIRR 544
Db 520 STARGRLGIRR 530
RESULT 8
ID AAB49706 standard; protein; 540 AA.
XX AAB49706;
AC AAB49706;
DT 04-APR-2001 (first entry)
DE Small round structured virus protein SEQ ID 7.
DE Small round structured virus; SRSV; food poisoning.
DE Small round structured virus.
DE WO200079280-A1.
DE 28-DEC-2000.
DE 22-JUN-2000; 2000WO-JP004095.
DE 22-JUN-1999; 99JP-00175928.
DE (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
DE (DENK-) DENKA SEIKEN KK.
DE Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
DE WPI; 2001-080848/09.
DE N-PSDB; AAF29147.
DE Kit for the detection and typing of small round-structured virus (SRSV)
DE strains for investigation of food poisoning outbreaks, contains
DE antibodies.
DE Claim 1; Page 54-57; 84pp; Japanese.
DE This invention relates to a kit for the detection and typing of small
DE round structured virus (SRSV) strains. The kit contains antibodies
DE directed against peptides represented in sequences AAB49700 - AAB49710,
DE which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
DE AAF20151 represent cDNA encoding the strain specific proteins. The kit is
DE used for detecting and typing strains of SRSV in order to prevent the
DE spread of infection and to examine the epidemiology of outbreaks
DE
DE Sequence 540 AA;
Query Match 42.3%; Score 1225; DB 4; Length 540;
Best Local Similarity 46.7%; Pred. No. 2.3e-100; Indels 36; Gaps 14;
Matches 261; Conservative 89; Mismatches 173;
Qy 1 MMWASKDAPTNMDGTSAGQLVPEANTAEPIISMFEVAGAAATAAGVNMIDPWIMNNY 60
Db 1 MKWASNDATPSNDGAAG---LVPEENN-EMALEPVGASLAAPVTGQTNIIDPWIRNF 56
Qy 61 VOAPGGEFTISPNTPGDIILDLQPHINPFLSHLAOMYNGVGNMKVYLLAGNAFTA 120
Db 57 VOAPNGEFTVSPRNSPGSILVNLELPELNLVLAHLARVNYAGGMEVQVNLAGNAFTA 116

Qy 121 GKIIISCIIPGFAAONISIAQATMFPHVIAIDVRVLEPIEVPLDVRNVLPH-NNDNAPT 179
Db 117 GKIIIFAAVPPYFPVENISPSQITMFPHVIIDVRLTEPVLDPMDVRLTFHFNQKDEPKM 176
Qy 180 RLVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVQOKTKPSPVNLPL 239
Db 177 RUVAMLYTPLASNG--SGDDVFTVSCRLTRPSPDFTLVLPPTVESKTKPFTLPVLT 234
Qy 240 NTLNSRVPISLIKMMVSRDQGVQONGRVTLDDGQLQGTTPTSASQLCKIRGSVFHAN 299
Db 235 GELNSRPFSLSIDEMVTSPNESIVVQPNQGRVTLDELLGTLTQACNICIRKVTGQV 294
Qy 300 GNGY-----NLTELDGSPYHAFES-PAPIGFPDL-GECDWHM-----EASPTTQFNTGD 347
Db 295 PSEQHMNLLEITNLNGTQFDPTDDVPAPLGVDPDFAGEVFGVLSQRNRCESNPANRAHDV 354
Qy 348 VIKQINVKOESAFAPHLGTIQAD--GLSDSVNTNMIAKLGWSPVSGHGRGDVDPWIP 405
Db 355 V-----ATYSDKYTPKGLGVQIGTWTNTDVENOPTKFTPIG-LNEVANGHR--FEQWTL 406
Qy 406 RYGSTLTEAAQLAPPYPPGGEAIVFFMSDFPIAHGTNGLSVPTIPQEFVTHFVNEQA 465
Db 407 RYSGALTLMNLAPAVAPLFGERLLFRSYVPLKGGFGNPAIDCSVQEWVQHFYQESA 466
Qy 466 PTRGEAALLHYLDPDTHRNIGBKLYPEGEMTCVPNSSGTPGOTPLPINGVFFVSVWSR 525
Db 467 PSLGDVALVRYNPDTRGVLFKALHKGGLTV--SSTSTGPPVVVPANGYFKFDSWVNOF 524
Qy 526 YQLKPVGTAGPACRLGIRR 544
Db 525 YSLAPMGITGN-----GRRR 538
RESULT 9
ID AAB49705 standard; protein; 548 AA.
XX AAB49705;
AC AAB49705;
DT 04-APR-2001 (first entry)
DE Small round structured virus protein SEQ ID 6.
DE Small round structured virus; SRSV; food poisoning.
DE Small round structured virus.
DE WO200079280-A1.
DE 28-DEC-2000.
DE 22-JUN-2000; 2000WO-JP004095.
DE 22-JUN-1999; 99JP-00175928.
DE (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
DE (DENK-) DENKA SEIKEN KK.
DE Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
DE WPI; 2001-080848/09.
DE N-PSDB; AAF29146.
DE Kit for the detection and typing of small round-structured virus (SRSV)
DE strains for investigation of food poisoning outbreaks, contains
DE antibodies.
DE Claim 1; Page 52-54; 84pp; Japanese.
DE This invention relates to a kit for the detection and typing of small
DE round structured virus (SRSV) strains. The kit contains antibodies
DE directed against peptides represented in sequences AAB49700 - AAB49710,
DE which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
DE AAF20151 represent cDNA encoding the strain specific proteins. The kit is
DE used for detecting and typing strains of SRSV in order to prevent the
DE spread of infection and to examine the epidemiology of outbreaks
DE
DE Sequence 540 AA;
Query Match 42.3%; Score 1225; DB 4; Length 540;
Best Local Similarity 46.7%; Pred. No. 2.3e-100; Indels 36; Gaps 14;
Matches 261; Conservative 89; Mismatches 173;
Qy 1 MMWASKDAPTNMDGTSAGQLVPEANTAEPIISMFEVAGAAATAAGVNMIDPWIMNNY 60
Db 1 MKWASNDATPSNDGAAG---LVPEENN-EMALEPVGASLAAPVTGQTNIIDPWIRNF 56
Qy 61 VOAPGGEFTISPNTPGDIILDLQPHINPFLSHLAOMYNGVGNMKVYLLAGNAFTA 120
Db 57 VOAPNGEFTVSPRNSPGSILVNLELPELNLVLAHLARVNYAGGMEVQVNLAGNAFTA 116

[illegible]

QY 121 KIIISCIIPPGFAAQNISIAQATMPPHVIAVRVLEPIEVLDPVRLFH-NNDNAPT 179
Db 117 KIIIFAAVPPNPTEGLSPQVTPPHIIVDVRLQLEPLILPDPVRNNFYHYNQSDSTI 176
QY 180 RLVCMLYTPLRASGSSGTDPPVIAGRVLTCPSPDFSLFLVPPNVEOKTKPSPVNLPL 239
Db 177 KLIAMLYTPLRA--NNAGDDVFTVSCRVLTRPSPDFILVPPVSRKTPFTVPIITV 234
QY 240 NTLNSRVPFLIKSMVSRDHQVQFQNGRVTLTGQGLQGTTPTSASQLCKIRGSVFPHAN 299
Db 235 EEMSNSRFPPIPLEKLYTGPSSAFVQPNQGRCTTDGVLGTTQLLSAVNICTPRGDVTHIA 294
QY 300 GONGY--NLTELDSGSPYHAFES-PAPIGFPPD-LGSCDWHMEASPTQNTGDVLIKQINVK 355
Db 295 GSHDYTNLASQWNSYNDTEIPAPLGTDPFVGKIQGML--TOTREDGSGTRAHKATVS 352
QY 356 QESA-FAPHLGTIQADGLSDVSNTNMIKLGWSPVSDG--HRGDVDPWIPRYGSTLT 412
Db 353 TGSVHTPKLGSVQVTTDNTNDFQTGQNTKFTPVGVIOQGNHQNPEQWVLPNTSGRTG 412
QY 413 EAAQLAPPIYPGFGAIVFFMSDPPIAHGTNGLSVPCTIPEFVTHFVNQAPTRGEAA 472
Db 413 HNVHLAPAVAPTFPCEQLLFFRSTMPGCGYPNNMLDCLLQEWVQHFCQEAAPQSDVA 472
QY 473 LLHYLDPTHNLGEBFKLYPBGFWMTCPVNSSGTGQTL--PINGVVFVSVWSRYQLKP 530
Db 473 LLRFVNPDTGRVLFECKLHSGYTVV---AHTGPHDLVIPPNGYFRFDSWVNOFTLAP 528
QY 531 VGTAGPACRLGIRRS 545
Db 529 MNGA-----GRRRA 538

RESULT 13

ID AAU91272 standard; protein; 548 AA.

AC AAU91272;

DT 18-JUN-2002 (first entry)

DE Norwalk virus associated polynucleotide #1.

XX Norwalk virus; monoclonal antibody; geno group I; geno group II;
immunological detection; food; viral infection.

OS Norwalk virus.

XX JP2002020399-A.

XX 23-JAN-2002.

XX 10-JUL-2000; 2000JP-00208151.

XX 10-JUL-2000; 2000JP-00208151.

XX (OSAP) OSAKA PREFECTURE.

PA (NISE-) NIPPON SEIBUTSU KAGAKU CENT KK.

PA (IATR) IATRION LAB INC.

XX WPI; 2002-287412/33.

XX A monoclonal antibody useful in the immunological detection and diagnosis
of Norwalk virus infection.

PS Disclosure; Page 12-13; 24pp; Japanese.

XX The invention describes a monoclonal antibody recognising Norwalk virus,
a capsid protein of Norwalk virus, or a common antigen epitope on the
capsid protein molecule of geno group I and geno group II. The antibody
is useful for immunological detection and quantitative analysis of
Norwalk virus in foods and the serum of infected patients. This sequence
represents a Norwalk virus associated protein described in the invention

XX Sequence 548 AA;
SQ Query Match
Best Local Similarity 40.0%; Score 1157; DB 5; Length 548;
Matches 254; Conservative 83; Mismatches 177; Indels 42; Gaps 15;
QY 2 MVASKDAPTNDGTSAGQOLVPEANTAEPIISMPEVAGAAATAATAGQVNMIDPWTMMNVV 61
Db 3 MASSRAAPSN-DGAAG--LVPEINN-EAMALDPVAGAAIAAPLTGQQNIIDPWTMMNVV 57
QY 62 QAPOGEFTISNNTPGDILOLQPHLNPFLSHLAQNYNGVGMKVKVLLAGNAFTAG 121
Db 58 QAPGGEFTVSRNPGEVLLLELPEINPYLAHLARMYNGYAGGFVQVVLGNAFTAG 117
QY 122 KIIISCIIPPGFAAQNISIAQATMPPHVIAVRVLEPIEVLDPVRLFH-NNDNAPT-MR 180
Db 118 KIIIFAAVPPNPIDNLSAQITMCPHVIVDVRLQLEPLILPDPVRNNFYHYNQSDSTLR 177
QY 181 LVCMLYTPLRASGSSGTDPPVIAGRVLTCPSPDFSLFLVPPNVEOKTKPSPVNLPLN 240
Db 178 LIAMLYTPLRA--NNSGDDVFTVSCRVLTRPSPDFSNFLVPPTVESKTKPFTLILIS 235
QY 241 TLSNSRVPFLIKSMVSRDHQVQFQNGRVTLTGQGLQGTTPTSASQLCKIRGSVFPHAN 300
Db 236 EEMSNSRFPPIESLHTSPNTENIVVQCQNGRVTLTGELMGTTLQLPSRICAFRGVLTSTRS 295
QY 301 G-----NGY---NLTELDSGSPYHAFES-PAPIGFPPD-LGSCDWHMEASPTTQ 342
Db 296 RASDAQDIATPRLFNYYWHVQDLNLTGTPYDPAEDIPGLTPTDPRGVK-----FGVASQ 350
QY 343 FNTGDVLIKQINVKQESA---FAPHLGTIQADGLSDVSNTNMIKLGWSPVSDG--HRG 397
Db 351 RNPDSITTRAHEAKVDTTAGRTPKLSLEISTESS-DFDQNPQTRF---TPVGIGVDNEA 406
QY 398 DVDPWIPRYGSTLTETAAQLAPPIYPGFGGAIUVFMSDFDIAHGTNGLSVPCTIPEBFV 457
Db 407 DFOQWLSLPDYSQGFTHNNMLAPAVAPNPPGSQLLFRRSQLESNGVLDCLVPQBW 466
QY 458 THFVNEQAPTRGEAALLHYLDPTHNLGEBFKLYPBGFWMTCPVNSSGTGPTPLPINGVVF 517
Db 467 QHFYQESAPAQTVQALVRYVNPDTGRVLFEAKLHKLGFMTIAKN--GDSPTVPPNGYFR 524
QY 518 FVSWVSRFYQLKPVGT 533
Db 525 FESWVNPFTLAPMGT 540

RESULT 14

AAU91272

ID AAB49709 standard; protein; 550 AA.

XX AAB49709;

AC AAB49709;

XX 04-APR-2001 (first entry)

XX Small round structured virus protein/SEQ ID 10.

XX Small round structured virus; SRSV; food poisoning.

XX Small round structured virus.

XX WO2000079280-A1.

XX 28-DEC-2000.

XX 22-JUN-2000; 2000WO-JP004095.

XX 22-JUN-1999; 99JP-00175928.

XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

XX (DENK-) DENKA SEIKEN KK.

XX		Small round structured virus protein SEQ ID 11.
DE	XX	Small round structured virus; SRSV; food poisoning.
KW	XX	Small round structured virus.
OS	XX	Small round structured virus.
XX		WQ200079280-A1.
PN	XX	
PD	XX	28-DEC-2000.
XX	XX	22-JUN-2000; 2000WO-JP004095.
Pf	XX	22-JUN-1999; 99JP-00175928.
PR	XX	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA	XX	(DENK-) DENKA SEIKEN KK.
XX	XX	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
PI	XX	WPI; 2001-080848/09.
DR	XX	N-PSDB; AAF29151.
DR	XX	Kit for the detection and typing of small round-structured virus (SRSV)
PT	XX	strains for investigation of food poisoning outbreaks, contains
PT	XX	antibodies.
XX		Claim 1; Page 64-66; 84pp; Japanese.
PS	XX	This invention relates to a kit for the detection and typing of small
CC	XX	round structured virus (SRSV) strains. The kit contains antibodies
CC	XX	directed against peptides represented in sequences AAB49700 - AAB49710,
CC	XX	which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC	XX	AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC	XX	used for detecting and typing strains of SRSV in order to prevent the
CC	XX	spread of infection and to examine the epidemiology of outbreaks
XX		Sequence 541 AA;
SQ		
		Query Match 35.0%; Score 1013; DB 4; Length 541;
		Best Local Similarity 40.5%; Pred. No. 2.1e-81;
		Matches 240; Conservative 79; Mismatches 154; Indels 120; Gaps 18;
OY	1	MMWAKDAPTNDGTSGAGQLVPREANTAEPISEMFVAGAATAATACGVNMIDPWNNY 60
Dd	1	MKGASNDAAAPSSDGAAG--LVPEINN-EVMPLFPVAGASLATPVGQQNIIDPWIRNF 56
OY	61	VQAPGGEFTISNNTPCGILFDLQLGHLNPFLSHLAQMNGVGNMVKVLLAGNAFTA 120
Dd	57	VQAPAGEFTVSPRNSPGSILLDLGLPOLNPLYLAHLMYNGHAGGMEVOIVLAGNAFTA 116
OY	121	GKIITSCITPPGFAAQNISIAQTAMFPHVIADVRLPEIEVPLEEDVRNVLFH-NNDNAPTM 179
Dd	117	GKIIIFAAIPGGPYENLSQITMCMPHVIDVRQLEPLLPMPIWNNFFHYNQGDPKL 176
OY	180	RIVCMLYTLPRASGSSGTHDPVIAGRVLTCSPDSFLFLVPPNVEOKTFPFSVPNLPL 239
Dd	177	RLVAMLYTFLRA---NNSGDVFETVSCRVLTRKSPDFEFTFLVPPVESKTQFALPILKI 234
OY	240	NTLSNRVPSLIKSMWVRDHQMQVQNGRVLTLDQLOGTTPTTSASOLCKIRGSVFHAN 299
Dd	235	SEMTNSRFPVPVDVVMYTARENQVVQPQNGRVTLTGELLGTLTPLLAVNICFKFGVIAKN 294
OY	300	GG-NGY----NLTLDGSPYHAFE-SPARIGEPDLEGCDWM-----EASPTTO---- 342
Dd	295	GDVRSYRMDEIINTDTGTDPTDTEBTPGIGSPDFQGILFGVASQRNKNEQNPAATEAHEA 354
OY	343	-FWTG--DVTKQINVKQESAFAPHL-----GTIQADGLSDSVSNWNMIK 384
Dd	355	IINTGDHICPFQISSSEIVLTSPNLRCTNPQLPQSLGRGTILI--RSDNGHCMDMVG- 411
OY	385	LGVVSPV-----SDHRGDVD-----PWVIPRYGSTLTEAAQLAPP 420

Db 412 ---TSPTTTPQWRRCRSGNCCSGHRYVPVVMNRVTWIVLSHKSGFSTSTRKLPQ 468
Qy 421 IYPEGFGAIVFFMSDFPIAHGTNGLSVFCTIPOEFVTHFVNEQAPTRGEAALLHYLDPD 480
Db 469 L-----NLRW-----LIRFINPD 482
Qy 481 THRLGEFKLYPEGFMTCVPNSSGTGPOTLPINGVFVSVWSRPYQLKPVGT 533
Db 483 TGRVLFEARLHKQGFITVA--HTGDNPIVMPPNNGYFRFEAWNQFYS LAPVGT 533

Search completed: June 1, 2004, 13:45:44
Job time : 48.8188 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 12.9849 Seconds
(without alignments)
2166.837 Million cell updates/sec

Title: US-09-926-799-1

Perfect score: 2896

Sequence: 1 MMASKDAPTNDTSGAGQ.....YQLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1950	67.3	530	4	US-08-486-049-3
2	285	9.8	626	4	US-09-590-020-7
3	272.5	9.4	669	4	US-09-617-594A-2
4	272	9.4	623	4	US-09-590-020-4
5	272	9.4	668	4	US-09-617-594A-4
6	269	9.3	623	4	US-09-590-020-2
7	262.5	9.1	622	4	US-09-590-020-6
8	149	5.1	2206	1	US-07-852-260-2
9	149	5.1	2206	3	US-08-461-503-2
10	149	5.1	2206	3	US-08-465-250-2
11	106	3.7	615	4	US-09-252-991A-26695
12	106	3.7	1302	4	US-09-423-890-2
13	106	3.7	1493	4	US-09-423-890-8
14	106	3.7	1593	4	US-08-628-829-4
15	106	3.7	2318	3	US-09-091-219-24
16	106	3.7	2318	4	US-09-660-541-24
17	103.5	3.6	1170	4	US-09-749-588-2
18	102.5	3.5	588	1	US-08-460-860-4
19	102.5	3.5	1210	4	US-09-963-137-197
20	102	3.5	1048	4	US-09-171-699-10
21	100	3.5	544	4	US-09-615-192A-349
22	100	3.5	1601	4	US-08-345-473E-40
23	98	3.4	1209	4	US-09-749-588-4
24	97	3.3	956	4	US-09-134-078-63
25	97	3.3	3443	2	US-08-416-603-2
26	97	3.3	4536	4	US-09-079-030-1
27	96	3.3	934	4	US-09-252-991A-25635

Sequence 10, Appl
Sequence 25, Appl
Sequence 25, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 15, Appl
Sequence 15, Appl
Sequence 9, Appl
Sequence 73, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 71, Appl

ALIGNMENTS

RESULT 1

US-08-486-049-3
; Sequence 3, Application US/08486049
; Patent No. 6572862

GENERAL INFORMATION:

APPLICANT: Estes, Mary K

APPLICANT: Jiang, Xi

APPLICANT: Graham, David Y

TITLE OF INVENTION: Methods and Reagents to Detect and

TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fulbright & Jaworski L.L.P.

STREET: 801 Pennsylvania Ave., N.W.

CITY: Washington, D.C.

STATE:

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/486,049

FILING DATE: June 7, 1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Davis, Peter

REGISTRATION NUMBER: 36,119

REFERENCE/DOCKET NUMBER: 311.023

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-662-0200

TELEFAX: 202-662-4643

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 530 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-486-049-3

Query Match 67.3%; Score 1950; DB 4; Length 530;

Best Local Similarity 67.0%; Pred. No. 2.4e-189;

Matches 366; Conservative 66; Mismatches 96; Indels 18; Gaps 7;

Qy 1 MMASKDAPTNDTSGAGQLVPEANTAEPISEPVAGATAAATAGOVNIDPWNNY 60

Db 1 MMASKDAPTNDTSGAGQLVPEANTAEPISEPVAGATAAATAGOVNIDPWNNY 60

126 SCIPPGF-AAQNIQAATMPHVIADRVLEPIEVLDPVNRVLFHNNNDNAPTMRILVCM 184
189 IVPPEGVDPVOSTBMLQ---YPHVLFPARQVEPVIFCLPDLRSLYLHMSDITDTLSLVM 245
185 LY---TFLRASGSSSGTDPFVIAGRVLTCPSPDFSLFVPPNVEQTKPFSPVNLPLN 240
246 VYNDLINFYANDTSSGC---IVT---VETKPGDPFKPHLLKPPG-----S 285
241 TILNSRVPS-II---KSMVSRDH-GQMVQFQNGRVTLTGQLOQTTPTSASQLCKIRGSV 295
286 MLTHGVSFSDLIKPSSSLWIGNRHWSIDTF-----IIRFV 322
296 FHANGGVN-----ITELDGGSPYHA-----PESPA-PIGFDPD---G 329
323 FOANRHFENQTAGMSTPRPISVTITEQNGAKLIGVATDYIVPGIDGWDPTTIPG 382
330 EC-----DWHM-----EASPTQFNTGDVIK-QINVK-----QESAPA 361
383 ELIPAGDYAITNGTGNIDITATGYDTADIIKNNTFRGMYICGSLQORAWGDKKISNTAP- 441
362 PHLGTITQADGLSDSVN-TNMI-----AKLGVSPVSD-----GHRGDVDPWVI 404
442 --ITATLTDGNNKINPCNTIDQSKIYVFQDAHVKKQAQTSDDTLALLGYTGIGEOAI- 498
405 PRYGSTLTEAAQLA--PPIYPGFGAEIVFFMSDFPIAHGTNGLSVPCPTIPQEFVTHFN 462
499 ---GSDRDRVRISTLPETGARG-GNHPIFYKNSIKLGYVIRSIDV-----FNS 543
463 EQAPTGEAALLHYLDP-----DTHRLGEBFKLYPEGFMTCPVNSSGTGPQTLPI 512
544 QILHTSRQLSLNHYLLPPDSFVYRIIDNSGWSFIDIGSDGF--SFVGSFGKLEPDL 601
513 NGVVFV 517
602 SASYM 606

RESULT 3
US-09-617-594A-2
; Sequence 2, Application US/09617594A
; Patent No. 6541458
; GENERAL INFORMATION:
; APPLICANT: Augonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V. V.
; FILE REFERENCE: 454313-3151.1
; CURRENT APPLICATION NUMBER: US/09/617,594A
; CURRENT FILING DATE: 2000-07-14
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-617-594A-2

Query Match 9.4%; Score 272.5; DB 4; Length 669;
Best Local Similarity 25.9%; Pred. No. 1.4e-18;
Matches 106; Conservative 54; Mismatches 134; Indels 115; Gaps 19;
13 DGTSGAGQLVPEANT-AEPISEMPVAGATAATAAGQVNMIDPWIMNMYVQAPQGEFTI- 70
126 DGDSSI---TTPEQTLVGGVIAEPSAQWATAAATGKSVDSW-----ESFSFH 174
71 -----SPNTPGDILFDLQGLPHLNPFLSHLAQMYNGVMKVKVLLAGNAFTAGKIII 125
175 TSVNWSSTETQKILFKQSLGLPLNPLYLHLSKLVVWMSGVSDVRFISGSGVFGKLA 234

61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGVMKVKVLLAGNAFTA 120
61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGVMKVKVLLAGNAFTA 120
121 GKIIISICPPGFAQNTSIAQAATMPHVIADRVLEPIEVLDPVNRVLFHNNNDNAPTMR 179
121 GKIIISICPPGFAQNTSIAQAATMPHVIADRVLEPIEVLDPVNRVLFHNNNDNAPTMR 180
180 RLVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFVPPNVEQTKPFSPVNLPL 239
181 RLVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFVPPNVEQTKPFSPVNLPL 238
240 NTLNSRVPSLIIKSMVSRDHGQMVQFQNGRVTLTGQLOQTTPTSASQLCKIRGSVFHAN 299
239 SSLNSRAPLPISSMGTSIPDNVQSVQFQNGRCTLGRLVGTTPVLSLHVAKIRGT-----S 294
300 CGNGYNLTDLGSPVHAPESAPITGPDGLGCDWHMEASPTQFNTGDVIKQINVKQESA 359
295 NGTVINLTDLGTPHPPEGAPITGPDGLGCDWHIN---MTQFGHSSQTYQDVDTTPDT 351
360 FAPHLGTITQADGLSDSVN-TNMIAKLGVSPVSDGHRGDVDPWVPIPRYGSTLTEAAQLAP 419
352 FVPHLGSQAANGIG-----SGNVGVLSWISPPSHSGSQVDLWKIPVYGGSSITEATHLAP 407
420 PIYPPGFGAEIVFFMSDFPIAHGTNGLSVPCPTIPQEFVTHFNQAPTRGEAALLHYLDP 479
408 SVYPPGFGAEIVFFMSKMP---GPCAYNLPCLLPQEIYSHLASEQAPTVGEAALLHYVDP 464
480 DTHRLGEBFKLYPEGFMTCPVNSSGTGPQTLPIVGVFVSVWSRFFYOLKPKVGTAGPA-C 538
465 DTGRNLGEFKAYPDGFLTCVPNGASSGPOQLPVGIVFVSVWSRFFYOLKPKVGTASSARG 524
539 RLGIIR 544
525 RLGLRR 530

RESULT 2
US-09-590-020-7
; Sequence 7, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4-1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-7

Query Match 9.8%; Score 285; DB 4; Length 626;
Best Local Similarity 23.8%; Pred. No. 6.9e-20;
Matches 144; Conservative 87; Mismatches 194; Indels 180; Gaps 32;
13 DGTSGAGQLVPEANT-AEPISEMPVAGATAATAAGQVNMIDPWIMNMYVQAPQGEFTI- 70
82 DGSITA---PEQTLVGGVIAEPSAQWATAAATGKSVDSW-----EAFSFFH 128
71 -----SPNTPGDILFDLQGLPHLNPFLSHLAQMYNGVMKVKVLLAGNAFTAGKIII 125
129 TSVNWSSTETQKILFKQSLGLPLNPLYLHLSKLVVWMSGVSDVRFISGSGVFGKLA 188


```
Qy 126 SCIPPGF-AAQNISIAQATMPHHVIADVRVLEIEVPLEVDRVNLFFHNDNAPTMRVLVCM 184
Db 235 IVPBPGVDPVQSTSMQ---YPHVLFARQVEPIFGISDPLRSTLYHLSMDTDTTSLVIM 291
Qy 185 LY-----TFLRASGSSGTDPPVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSVPLN 240
Db 292 VYNDLIMPYANDSSGC---IVT--VETKPGDPKFHLKPPG-----S 331
Qy 241 TLSNRVPS--LIKSMV---SRDHGMVQVQNGRVTLTGLOQGTPTTSASOLCKIRGSV 295
Db 332 MLTHGSPSLDLPKSSSLWGNRYMSDITDF-----VIRPFV 368
Qy 296 PHANGGYN-----LTLDGSPYHA-----PESPA-PIGFPD--LGE 330
Db 369 FOARHDFNOETAGWSTPRFRPITITISSENGSKLGTGVATDVIIVPGIPDGPDTTIGE 428
Qy 331 CDHWEASPTTQFNTGDIKQINVKQESAFAPHLGTIQADGLSDVSUNT 379
Db 429 -----ELTPAGDYSI-----TNGSGNDATANAYDSADVINT 461

RESULT 4
US-09-590-020-4
; Sequence 4, Application US/09590020
; Patent No. 635246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilinis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-4

Query Match 9.4%; Score 272; DB 4; Length 623;
Best Local Similarity 22.2%; Pred. No. 1.4e-18;
Matches 133; Conservative 84; Mismatches 209; Indels 174; Gaps 27;

Qy 7 DAPT-NMDGTSGAGQLVPEANTAPFISMEPVAGAAATAAGQVNMIDPMTMNNYVOAPQ 65
Db 72 DLPLRLEGDDGSIITPEQGTMGVGVIAEPSPAQMSAAADMATGKSDSEW-----E 122
Qy 66 GEFTI-----SPNNTPGDILFDLQGLHNPFLSHLAQMYNGVGNMKVVLGNAFT 119
Db 123 AFFSFTSVNMSTSETQCKILFKQSLGFLNPNYLTHLAKLYVAMSGSVSEVRFSGSGVF 182
Qy 120 AGKIIICIPPGF-AAQNISIAQATMPHHVIADVRVLEIEVPLEVDRVNLFFHNDNAPT 178
Db 183 GKGKAAIIVVPGIEPVQSTSMQ---YPHVLFARQVEPIFAIPDLKSNLYHLSMDTDT 239
Qy 179 MRLVCMLY-----TFLRASGSSGTDPPVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSV 234
Db 240 TSLVIMVYNDLIMPYANDSSGC---IVT--VETKPGDPKFHLKPPG-----284
Qy 235 PNLPLNTLSNRVPS-LI---KSMVSRDHGMVQ-----FQ-NGRVTLTGLOQ---279
Db 285 -----SMLTHGSPSLDLPKSSSLWGNRYMSDITDFIIRPFVQANRHFDFNOETAGWS 339
Qy 280 -----TTPTTSASOLCKI-----RGSVPHANGGYN 305
Db 340 TPRFRPITITVSESNMKSGLGVATDVIIVPGIPDGPDTTIPBQLTAPAGIYSITASNGTD 399
Qy 306 LTLDGSPYHAFESPA-PIGFPDLGECD-----WHMEASPTTQF-----NT 345
```

```
Db 400 ITTAAG--YDAAEITVNTTFFKSMYICSLQRAWGDKKISNTAFITTAVRKGNSEPSNT 457
Qy 346 GDTVIKQINVKQESAFAPHLGTIQADGLSDVSVNTNMIKLGWSPVSDGHRGDDVP---- 401
Db 458 IDMTKL-VVQDA-----HVG-----EVQTSITLALLGYTGIGEAIGSDRDKVRI 505
Qy 402 WVIPRYGSTITEAAQAAPPYPPGFGBAIVFFMSDFPIAHGTNGLSVPTTIPQBFVTHFV 461
Db 506 SVLPETGAR-----GGNHPIFYKNSIKLGVVIRSIDV-----FN 539
Qy 462 NEQAPTGEAALLHYLDP-----DTHRNLGEFKLYPEGEFTMCPVNSSSGTGPTLP 511
Db 540 SQLHTSRQLSNLNYLLPPDSFAVYRILDSNGSDFDIDTDF-----SFVGSVNL 592

RESULT 5
US-09-617-594A-4
; Sequence 4, Application US/09617594A
; Patent No. 6541458
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.1
; CURRENT APPLICATION NUMBER: US/09/617,594A
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-617-594A-4

Query Match 9.4%; Score 272; DB 4; Length 668;
Best Local Similarity 26.7%; Pred. No. 1.6e-18;
Matches 107; Conservative 46; Mismatches 134; Indels 114; Gaps 20;

Qy 23 PEANT-----AEP-ISMEPVAGAAATAAGQVNMIDPMTMNNYVOAPQGEFTI----- 70
Db 133 PEQTLVGGVIAEPNAQMSAVADVATGKSDSE-----W-----EAFSPHTSVN 177
Qy 71 -SPNNTPGDILFDLQGLHNPFLSHLAQMYNGVGNMKVVLGNAFTAGKIIISCIP 129
Db 178 WSTSETQCKILFKQSLGFLNPNYLTHLAKLYVAMSGSVSEVRFSGSGVFGKLAIIVVP 237
Qy 130 PGF-AAQNISIAQATMPHHVIADVRVLEIEVPLEVDRVNLFFHNDNAPTMRVLCMLY-- 186
Db 238 PGIDPVQSTSMQ---YPHVLFARQVEPIFTIPDLKSNLYHLSMDTDTTSLVIMVYND 294
Qy 187 --TFLRASGSSGTDPPVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSVNLPLNTLSN 244
Db 295 LINPYANDSSGC---IVT--VETKPGDPKFHLKPPG-----SMLTH 334
Qy 245 SRVPS-LI---KSMVSRDH--GMVQVQNGRVTLTGLOQGTPTTSASOLCKIRGSVFHAN 299
Db 335 GSIPSDLPKSSSLWGNRYMSDITDF-----VIKPFVFOAN 371
Qy 300 GGNGYN-----LTLDGSPYH---AFESPAPIGFPDLGECDDHMEASP 339
Db 372 RHDFNOETAGWSTPRFRPITITVSEKSGSKLGTGVATDTSIVP-GIPD-----GWPDTTIP 426
Qy 340 TTQFNTGDIKQINVKQESAFAPHLGTIQADGLSDVSUNT 380
Db 427 EKLTPAGDYAITNGNNDITTA-----ADYDGAIIKNTN 462
```

```

US-09-590-020-2
; Sequence 2, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09590.020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
; ORGANISM: Feline calicivirus
US-09-590-020-2

Query Match          9.3%; Score 269; DB 4; Length 623;
Best Local Similarity 29.8%; Pred. No. 2.9e-18;
Matches 76; Conservative 41; Mismatches 94; Indels 44; Gaps 9;

QY 7 DAPT-NMDGTSGAGQLVPEANTAEPISEPVAGATAAATAAGQVNMIDPWIMNNYVQAPQ 65
DB 72 DLPLFRLGGDGSITTPGCTMVGVIAPESQAQMSAAADMATGKSVDSW-----E 122
QY 66 GEFTI-----SPNTPGDLFDLQGLPHLNPFLSHLAQMYNGVGMKVKVLLAGNAFT 119
DB 123 AFPSFHTSNWSTSETQKILFKQSLGPLNLYSLAKLYVALAGSVEVRFSGSGVF 182
QY 120 AGKIISCIPPGF-AAQNISIAQATMPHVIADRVLEPIEVLDRNVLPHNNDNAPT 178
DB 183 GKGKLAIVVPPGIEPVQSTMLQ---YPHVLFDAQVEVIFAIPDLRSNLHLMSDITD 239
QY 179 MRLVCMXY-----TPLRASGSSGTDPPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSV 234
DB 240 TSLVIMVYNDLINPYANDTNSSG---IVT--VETKPGDPFKHLKPPG----- 284
QY 235 PNLPLNTLSNSRVS-LI---KSMVSRDHQMVQ-----FQ-NGRVTLDGQLQ-- 279
DB 285 -----SMLTHGSPVPSDLIPKSSSLWIGNRHSWSDITDFIIRPFVQANRHFDFNOETAGWS 339
QY 280 -----TTPTSASQLCKI-----RGSVFHANGGNGYN 305
DB 340 TFRPRFITVSESNMKGIGVATDYIVPGIDGWPDTTPEQLTPAGIYSITASNGTD 399
QY 306 LTELDGPYHAFESPA-PIGFPDLGEC-----WHMEASPTTQF-----NT 345
DB 400 ITTAAG--YDAETIVNTTNFKSMVICGSLQRAWDKKISNTAFITTAVRKGNSEPTNT 457
QY 346 GDVIKQINVKQSAFAPHLGTTQADGLSDVSNTNMIAKLGWVSPVSDGHRGDVDPWV-I 404
DB 458 IDMTKLIV-VYQDA-----HVG-----EVQSDITLALLGYTGIGBEAIGSDRDKVRI 505
QY 405 PRVGSITLTAQAQLAPYPPGGEAIVFPMSDFPIAHGTNGLSVPCITPQEPVTHFVNEQ 464
DB 506 SVLGETGARG-----GNHPIFYKNSIKLGVIIRSIVD-----FNSQI 542
QY 465 APTRGEAALHY-LIDPDTHR-----NLGEFKLYPEGPMTCVPSNPGTGPQLP 511
DB 543 LHSSRQLSLNLYLPPDSFAVIRLDSNGSWFDIGIDTFEPDGF-----SFVGVSNLP 595

RESULT 6
US-09-590-020-2
; Sequence 2, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09590.020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Feline calicivirus
; ORGANISM: Feline calicivirus
US-09-590-020-6

Query Match          9.1%; Score 262.5; DB 4; Length 622;
Best Local Similarity 22.3%; Pred. No. 1.3e-17;

```

```

Matches 134; Conservative 82; Mismatches 213; Indels 171; Gaps 28;

QY 7 DAPT-NMDGTSGAGQLVPEANTAEPISEPVAGATAAATAAGQVNMIDPWIMNNYVQAPQ 65
DB 72 DLPLFRLGGDGSITTPGCTMVGVIAPESQAQMSAAADMATGKSVDSW-----E 122
QY 66 GEFTI-----SPNTPGDLFDLQGLPHLNPFLSHLAQMYNGVGMKVKVLLAGNAFT 119
DB 123 AFPSFHTSNWSTSETQKILFKQSLGPLNLYSLAKLYVALAGSVEVRFSGSGVF 182
QY 120 AGKIISCIPPGF-AAQNISIAQATMPHVIADRVLEPIEVLDRNVLPHNNDNAPT 178
DB 183 GKGKLAIVVPPGIEPVQSTMLQ---YPHVLFDAQVEVIFAIPDLRSNLHLMSDITD 239
QY 179 MRLVCMXY-----TPLRASGSSGTDPPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSV 234
DB 240 TSLVIMVYNDLINPYANDTNSSG---IVT--VETKPGDPFKHLKPPG----- 284
QY 235 PNLPLNTLSNSRVS-LI---KSMVSRDHQMVQ-----FQ-NGRVTLDGQLQ-- 279
DB 285 -----SMLTHGSPVPSDLIPKSSSLWIGNRHSWSDITDFIIRPFVQANRHFDFNOETAGWS 339
QY 280 -----TTPTSASQLCKI-----RGSVFHANGGNGYN 305
DB 340 TFRPRFITVSESNMKGIGVATDYIVPGIDGWPDTTPEQLTPAGIYSITASNGTD 399
QY 306 LTELDGPYHAFESPA-PIGFPDLGEC-----WHMEASPTTQF-----NT 345
DB 400 ITTAAG--YDAETIVNTTNFKSMVICGSLQRAWDKKISNTAFITTAVRKGNSEPTNT 457
QY 346 GDVIKQINVKQSAFAPHLGTTQADGLSDVSNTNMIAKLGWVSPVSDGHRGDVDPWV-I 404
DB 458 IDMTKLIV-VYQDA-----HVG-----EVQSDITLALLGYTGIGBEAIGSDRDKVRI 505
QY 405 PRVGSITLTAQAQLAPYPPGGEAIVFPMSDFPIAHGTNGLSVPCITPQEPVTHFVNEQ 464
DB 506 SVLGETGARG-----GNHPIFYKNSIKLGVIIRSIVD-----FNSQI 542
QY 465 APTRGEAALHY-LIDPDTHR-----NLGEFKLYPEGPMTCVPSNPGTGPQLP 511
DB 543 LHSSRQLSLNLYLPPDSFAVIRLDSNGSWFDIGIDTFEPDGF-----SFVGVSNLP 595

RESULT 8
US-07-852-260-2
; Sequence 2, Application US/07852260
; Patent No. 5525715
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/852,260
; FILING DATE: 19920619
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678

```

```
;
; REFERENCE/DOCKET NUMBER: 36607-B-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-852-260-2

Query Match      5.1%; Score 149; DB 1; Length 2206;
Best Local Similarity 20.3%; Pred. No. 3.8e-05;
Matches 97; Conservative 57; Mismatches 196; Indels 128; Gaps 19;

QY 15 TSGAGQLVPEANTAEPISEMEPVAGAAATAAGOV-NM-----IDPWIMNNYVQAPQG-- 66
Db 347 TPGSNQYLTSNHSQPCAI-PEFDVTPPIDIPGEVKNMELAEIDTMIPLNLESTKRTM 405
QY 67 ---BFTISPNNTPGDILFDLQGLHNPFLSH-----LAQMYNGMVGNMKVLLAGNAF 118
Db 406 DMYRVTLSDSADLSQPICLSLSPAFDRLSHTMLGEVLNYTHWAGSLKFTFLFCGSM 465
QY 119 TAGKIIISCIPPPGAAQNISIAQATMPHPHVIADVRVLEPIEVPLEDVRNVLFH--NNDNA 176
Db 466 ATGKILVAYAPPG-AQPPTSKEAMLGTHVLDGLQSSCTMVVPWISNVTYRTQDSF 524
QY 177 PTMLVCMLY-----TPLRASGSSGTDPPFVIAGRVLTCPSPDPSFLP-----V 221
Db 525 TEGGYISMFYQTRIVWPLSTPKMS-----MLGFVSAC--NDFSVRLLRDTTHISQSAL 576
QY 222 PPNVEQTK-----PFSVPNLP-----LNTLSNRVPSLIKSM 255
Db 577 PQGIEDLTSEVAQALTLSPKQDSDPTKASGPAHSKEVPALTAVETGATNPLASDT 636
QY 256 VSRDHGMVQ-----FQNGRVTLDGQLQGTTPTSASQLCKIRGVSFVH----- 297
Db 637 VQTRH--VVQRSSRSESTIESFFARGACVAIIEDNEQPTTRAQKLPAMWRITYKDTVOL 694
QY 298 -----ANGGNYNLTELDGSPYHAFESPAPIGFDPDLGECDW 333
Db 695 RRKLEFFTSRFDMEFTFVVVTANFTNANNGHALNQVYQIMYPFGAPTPKSWDDY---TW 751
QY 334 HMEASPTTQNTGDVIKQINVKQESSAFAPHLGTTQA-----DGLSDVSVNTNIAKLG 386
Db 752 QTSNPSIFITYGAAPARISV-----PYVGLANAYSHFYDGFPAKPLKTDANDQIG 802

RESULT 9
US-08-461-503-2
; Sequence 2, Application US/08461503
; Patent No. 5834302
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES
; TITLE OF INVENTION: FROM CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461.503
; FILING DATE: 5-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-D-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-503-2

Query Match      5.1%; Score 149; DB 2; Length 2206;
Best Local Similarity 20.3%; Pred. No. 3.8e-05;
Matches 97; Conservative 57; Mismatches 196; Indels 128; Gaps 19;

QY 15 TSGAGQLVPEANTAEPISEMEPVAGAAATAAGOV-NM-----IDPWIMNNYVQAPQG-- 66
Db 347 TPGSNQYLTSNHSQPCAI-PEFDVTPPIDIPGEVKNMELAEIDTMIPLNLESTKRTM 405
QY 67 ---BFTISPNNTPGDILFDLQGLHNPFLSH-----LAQMYNGMVGNMKVLLAGNAF 118
Db 406 DMYRVTLSDSADLSQPICLSLSPAFDRLSHTMLGEVLNYTHWAGSLKFTFLFCGSM 465
QY 119 TAGKIIISCIPPPGAAQNISIAQATMPHPHVIADVRVLEPIEVPLEDVRNVLFH--NNDNA 176
Db 466 ATGKILVAYAPPG-AQPPTSKEAMLGTHVLDGLQSSCTMVVPWISNVTYRTQDSF 524
QY 177 PTMLVCMLY-----TPLRASGSSGTDPPFVIAGRVLTCPSPDPSFLP-----V 221
Db 525 TEGGYISMFYQTRIVWPLSTPKMS-----MLGFVSAC--NDFSVRLLRDTTHISQSAL 576
QY 222 PPNVEQTK-----PFSVPNLP-----LNTLSNRVPSLIKSM 255
Db 577 PQGIEDLTSEVAQALTLSPKQDSDPTKASGPAHSKEVPALTAVETGATNPLASDT 636
QY 256 VSRDHGMVQ-----FQNGRVTLDGQLQGTTPTSASQLCKIRGVSFVH----- 297
Db 637 VQTRH--VVQRSSRSESTIESFFARGACVAIIEDNEQPTTRAQKLPAMWRITYKDTVOL 694
QY 298 -----ANGGNYNLTELDGSPYHAFESPAPIGFDPDLGECDW 333
Db 695 RRKLEFFTSRFDMEFTFVVVTANFTNANNGHALNQVYQIMYPFGAPTPKSWDDY---TW 751
QY 334 HMEASPTTQNTGDVIKQINVKQESSAFAPHLGTTQA-----DGLSDVSVNTNIAKLG 386
Db 752 QTSNPSIFITYGAAPARISV-----PYVGLANAYSHFYDGFPAKPLKTDANDQIG 802

RESULT 10
US-08-465-250-2
; Sequence 2, Application US/08465250
; Patent No. 6136570
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
```

us-09-926-799-1-rai

Wed Jun 2 09:13:23 2004

```
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release 1.30
; CURRENT APPLICATION DATA: US/08/465,250
; APPLICATION NUMBER: US/08/465,250
; FILING DATE: 6-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-E-PCT-US
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-465-250-2

Query Match 5.1%; Score 149; DB 3; Length 2206;
Best Local Similarity 20.3%; Pred. No. 3.8e-05;
Matches 97; Conservative 57; Mismatches 196; Indels 128; Gaps 19;

QY 15 TSGAGQLVPEANTAEPISEMPVAGATAAATAGOV-NM-----IDPWNNVVOAPQG-- 66
DB 347 TPGSNQYLSDNHQSPCAI-PFEDVTPRIDPCEVKMMELAEIDTIDPLNESTKRTM 405
QY 67 ----EFTISPNTPGDILDLQLGPHLNPELH-----LAQYNGVGVNKKVLLAGNAF 118
DB 406 DMVRVTLSDSADLSQPICLSLSPAPDRLSHMTLMGEVLNYTHWAGSLKFTFLFCGSM 465
QY 119 TAGKIIISCIIPGFAAQNISIAQATMFPFVIADRVLEPIEVLEPVDNRVLFH--NNDA 176
DB 466 ATGKILVAVAPPG-AQPTSRKEAMLGTHVINDLGQSSCTMVVFWISNVTYRQTQDSF 524
QY 177 PTMRLVCMLY-----TPLRAGSSSGTDPFVIAGRVLTCPSPDESFLFL-----V 221
DB 525 TEGGYISMFTQTRIVVPLSTPKSMS-----MLGFVSAC--NDFSVLLRDLTHISQSAL 576
QY 222 PNVEQKTK-----PFSVPNLP-----LNTLSNRVPSLIKSM 255
DB 577 PQGIEDLTSEVAGALTSLSPKQDLSLPDTKASGPAHSKEVPALTAETGATNPLAPSDT 636
QY 256 VSRDHGQMVQ-----FQNGRVTLDDQLGTTPTTSASQLCKIRGSVPH----- 297
DB 637 VQTRH--VVQRRSRSESTIESFFARGACVAIIEVDNEQPTTTRAKLFAWMRIYTKDTVOL 694
QY 298 -----ANGGVNLTDLGSPYHAFESPAPIGFPDLDGECDW 333
DB 695 RRKLEFFTSRDMFEFTVVVANTNNGHALNQVQIMYIPPGAPTPKSWDY--TW 751
QY 334 HMEASPTTQFTGTVIKQINVKQESAFAPHLGTIOA-----DGLSDVSVNTNMIKLG 386
DB 752 QTSNPSFIYTYGAAPARISV-----PYVGLANAYSHFYDGFVKPLKTDANDQIG 802

RESULT 11
US-09-252-991A-26695
; Sequence 26695, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
```

```
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26695
; LENGTH: 615
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-26695

Query Match 3.7%; Score 106; DB 4; Length 615;
Best Local Similarity 22.0%; Pred. No. 0.11;
Matches 107; Conservative 56; Mismatches 181; Indels 142; Gaps 24;

QY 4 ASKDAPTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGOVNMID 53
DB 133 SSSTAP-----RPGPGVLPVATPSSPPALGSPVGGVIVPEPPNSPPAPKSPVGGVVPVE 187
QY 54 PWIMNYVOAPQ-----GEFTISPNT-----PGDILDLQLGPHLNPFSLHQAQMYNG 102
DB 188 P---PNSPPAPKSPVGGVVPVEPPNSPPARRSVPGEV-----PVAPSSSTAPR--- 234
QY 103 WVGNNMKVLLAGNAFTAGKIIISCIIPGFAAQNISIAQATMFPFVIADRVLEPIEVPL 162
DB 235 -PGPGVLPVATPNSSTAPRSILGRVLP-VATPNSPPAPKSPVGGV-----PVEPP- 284
QY 163 EDVRNVLFHNNDNAPTMRLVCMLYTPLRASGS-----SSGTDPFVIAG 205
DB 285 -----NSPPAPKSPVGGVVPVEPPNSPPAPKSPVGGVVPVAPNSSTAPRSILG 333
QY 206 RVLTCPSPDFSLFVPPNVNVEOKTFPSVFN--LPLNTLSNRVPSLIKSMVMVRDHGOM 263
DB 334 RVLVPVATPS-----SPPALG-----SVPGGVIPVESPSNSTAP---RSIL----- 370
QY 264 VQFQNGRVTLDDQLGTTPTTSASQLCKIRGSVPHANGN-----GYNLTELDGSP 313
DB 371 -----GRV-----LPVATPSSPPALGSPVGGVIVPEPPNSPPAPKSPVGGVVPVEPPNSP 420
QY 314 YHAFESPAPIGFPDLGCDWHMEASPTTQNTGDVIKQINVKQESAFAPHLGTIOADGLS 373
DB 421 -----PAPKSPV--GEV--VPVAPNSSTAPRSILGRVLPVATPSSPPALGSPVGGVIP 470
QY 374 DVSVNTNMIK--LGMVSPV-----SDGHRGDVDPFWIPRYGSTLTAAQLAP---PIYP 423
DB 471 VESPSNSTAPRSILGRVLPVATPSSPPALGSPVGGVIVPEPPNSSTAPRLDPEGVVPVAP 530
QY 424 PGFGEA 429
DB 531 PSSSTA 536

RESULT 12
US-09-423-890-2
; Sequence 2, Application US/09423890
; Patent No. 6312934
; GENERAL INFORMATION:
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION
; TITLE OF INVENTION: HUMAN MEK PROTEIN AND NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: CPI-085CPPC
; CURRENT APPLICATION NUMBER: US/09/423,890
; CURRENT FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: USN 60/078,153
; PRIOR FILING DATE: 1998-03-16
; PRIOR APPLICATION NUMBER: USN 60/099,165
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1302
```


Db 867 EFYPHIVSTVDSQAEPEVIRYKLLSLTALQSIDNSHVMGKLSRRIY----- 916
Qy 197 GTDFPVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLNTLSNRVPSLIKSMVMV 256
Db 917 -----LSARWVTAVPAVFSKLVTM-----LNASGTHFTMRRLMA 954
Qy 257 SRDHGMVQ-FQNG-RVTLDG---QLQTTPTSQCLCKIRGSVFHANGNGYNLTELD- 310
Db 955 LADEVIAEIVQLGVEDVDGHQDSLQAVPTSCLENSLEHTVHREKTKGKLSATRLSA 1014
Qy 311 -----GSPYHAFESPAPGFDLGECDWHMEAS 338
Db 1015 SSEDISRLAGVSVGLPSTTTEQKPAVQTKGRPHSQCLNSPLSHAQL-----MPPA 1068
Qy 339 PTTQNTGDNVQINVKQESAFAP 362
Db 1069 PSAPCSAPSVPDISKRRPQAFVP 1092

RESULT 15
US-09-091-219-24
; Sequence 24, Application US/09091219
; Patent No. 6171592
; GENERAL INFORMATION:
; APPLICANT: STUDDERT, Michael J.
; APPLICANT: CRABB, Brendan S.
; APPLICANT: FENG, Li
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
; FILE REFERENCE: 040268/0151
; CURRENT APPLICATION NUMBER: US/09/091,219
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: PCT/AU96/00815
; EARLIER FILING DATE: 1996-12-18
; EARLIER APPLICATION NUMBER: AU PN7201
; EARLIER FILING DATE: 1995-12-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 2318
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
US-09-091-219-24

Query Match 3.7%; Score 106; DB 3; Length 2318;
Best Local Similarity 22.3%; Pred. No. 0.96;
Matches 63; Conservative 31; Mismatches 113; Indels 76; Gaps 13;
Qy 5 SKDPTNMDCTSGAGQLVPEANTAEPISME-----PVAGATAAATAGQVNMIDPWIMNYY 60
Db 270 SKLASSAFSLFGA--LLADKTEETLLEDRILTTTRNGHTTSTTQSSGV-----TY 320
Qy 61 VQAPQGEFTTSPNNTPGDI-----LFDLQLGP-----HL-----NPF 92
Db 321 GYATAEDFVGSPNTSGLETRVQAEFPFKTHLFDWTSDFSGRCHLELFTDHHKGVYGS 380
Qy 93 LSHLAQMYNGVGNMKYKLLAGNAFTAGKIIISCIPPGFAQNIISIAQATMEPPHVIADV 152
Db 381 TDSYAYVRNGW-----DVEVTAVNGQFNGGCLLVAMVPELYSIQKRELYQLTLPHQFINP 436
Qy 153 R--VLEPIEVPLEDRVN-----LFHNDNAPTMRLLVCMLY--TPLRA 191
Db 437 RTNNTAHTVFPVGVNRYDQYKHKPTLVVMVVAPLTVNTGAPQIKVYANIAPTNVHV 496
Qy 192 SG-----SSSGTDPFVIA-----GRVLTCP-----SPDFSFLFLVPPN 224
Db 497 AGEFPFSKEGIPFVACSDGYGGLVTTDPKTADPVYGVNFPNPN 539

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 34.8398 Seconds
(without alignments)
4368.312 Million cell updates/sec

Title: US-09-926-799-1
Perfect score: 2896
Sequence: 1 MMASKDAPTNMDGTSGAQ.....YQLKPVGTAGPACKLGIIRS 545

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071 .

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pdb.*

2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pdb.*

3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pdb.*

4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pdb.*

5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pdb.*

6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pdb.*

7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pdb.*

8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pdb.*

9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pdb.*

10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pdb.*

11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pdb.*

12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pdb.*

13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pdb.*

14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pdb.*

15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pdb.*

16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pdb.*

17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pdb.*

18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pdb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1950	67.3	530	14	US-10-314-739-3
2	276	9.5	547	12	US-10-670-695-4
3	276	9.5	671	12	US-10-670-695-2
4	272.5	9.4	668	14	US-10-209-507-2
5	272	9.4	668	14	US-10-209-507-4
6	122.5	4.2	6310	12	US-10-282-122A-67793
7	122	4.2	913	15	US-10-369-493-2294
8	122	4.2	3263	12	US-10-282-122A-77663
9	114.5	4.0	1765	12	US-10-282-122A-48055
10	113	3.9	724	14	US-10-156-761-9262
11	113	3.9	1555	14	US-10-128-714-3298
12	113	3.9	1832	14	US-10-128-714-8298
13	112.5	3.9	2914	15	US-10-093-463-82
14	112	3.9	3930	12	US-10-282-122A-46817
15	111	3.8	1046	14	US-10-224-999A-3480

16	108.5	3.7	1194	12	US-10-282-122A-46163	Sequence 46163, A
17	108	3.7	5935	14	US-10-243-243A-8	Sequence 8, Appli
18	107.5	3.7	585	15	US-10-369-493-15469	Sequence 15469, A
19	107.5	3.7	2703	12	US-10-282-122A-66108	Sequence 66108, A
20	107	3.7	5877	14	US-10-142-515-11	Sequence 11, Appl
21	106	3.7	507	12	US-10-424-599-185988	Sequence 185988, A
22	106	3.7	510	12	US-10-425-114-72837	Sequence 72837, A
23	106	3.7	1302	13	US-10-000-864-2	Sequence 2, Appli
24	106	3.7	1493	9	US-09-858-754-3	Sequence 3, Appli
25	106	3.7	1493	13	US-10-000-864-8	Sequence 8, Appli
26	105.5	3.6	818	14	US-10-174-677-86	Sequence 86, Appl
27	105.5	3.6	932	14	US-10-174-677-87	Sequence 87, Appl
28	104.5	3.6	946	11	US-09-764-875-638	Sequence 638, App
29	104	3.6	435	15	US-10-609-775-24	Sequence 24, Appl
30	104	3.6	800	15	US-10-108-260A-2490	Sequence 2490, Ap
31	104	3.6	1052	12	US-09-905-211-2	Sequence 2, Appli
32	103.5	3.6	1170	13	US-10-135-687-2	Sequence 2, Appli
33	103.5	3.6	1210	9	US-09-860-352A-2	Sequence 2, Appli
34	103.5	3.6	1210	15	US-10-410-764-89	Sequence 89, Appl
35	103.5	3.6	1210	15	US-10-258-106-1	Sequence 1, Appli
36	102.5	3.5	869	12	US-10-087-192-1683	Sequence 1683, Ap
37	102.5	3.5	1210	12	US-09-963-131-197	Sequence 197, App
38	102	3.5	1048	14	US-10-223-538-10	Sequence 10, Appl
39	101.5	3.5	1021	16	US-10-443-101-2	Sequence 2, Appli
40	101	3.5	616	14	US-10-156-761-10270	Sequence 10270, A
41	100.5	3.5	1070	9	US-09-735-367B-6	Sequence 6, Appli
42	100	3.5	544	14	US-10-174-693-349	Sequence 349, App
43	100	3.5	828	9	US-09-738-626-5038	Sequence 5038, Ap
44	100	3.5	1283	15	US-10-369-493-22781	Sequence 22781, A
45	100	3.5	1589	15	US-10-369-493-5611	Sequence 5611, Ap

ALIGNMENTS

RESULT 1
US-10-314-739-3
Sequence 3, Application US/10314739
Publication No. US20030129588A1

GENERAL INFORMATION:
APPLICANT: Estes, Mary K
Jiang, Xi

Graham, David Y

TITLE OF INVENTION: Methods and Reagents to Detect and

Characterize No. US20030129588A1walk and Related Viruses

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fulbright & Jaworski L.L.P.

STREET: 801 Pennsylvania Ave., N.W.

CITY: Washington, D.C.

STATE: <Unknown>

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/314,739

FILING DATE: 09-Dec-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/486,049

FILING DATE: June 7, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Davis, Peter

REGISTRATION NUMBER: 36,119

REFERENCE/DOCKET NUMBER: 311.023

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-662-0200

TELEFAX: 202-662-4643

TELEX: <Unknown>

```

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 530 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-314-739-3

Query Match      67.3%; Score 1950; DB 14; Length 530;
Best Local Similarity 67.0%; Pred. No. 6.1e-180; Indels 18; Gaps 7;
Matches 366; Conservative 66; Mismatches 96;

QY 1 MMASKDPTNMDGTSAGQQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNMY 60
DB 1 MMASKDPTNMDGTSAGQQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNMY 60
QY 61 VQAPQGEFTISNNTPGDIILDLQGLPHNPFSLHQAQMYNGVGNMKVYKVLNAGNFTA 120
DB 61 VQAPQGEFTISNNTPGDIILDLQGLPHNPFSLHQAQMYNGVGNMKVYKVLNAGNFTA 120
QY 121 GKIIISCIPPGAAQNIATAQTMFPHVIADVRVLEPIEVLDPVRLVFNHND-NAPTM 179
DB 121 GKIIISCIPPGAAQNIATAQTMFPHVIADVRVLEPIEVLDPVRLVFNHND-NAPTM 179
QY 121 GKIIIVSCIPPGSHNLTIAQATLFPFHVADVRVLEPIEVLDPVRLVFNHNDNQTM 180
DB 121 GKIIIVSCIPPGSHNLTIAQATLFPFHVADVRVLEPIEVLDPVRLVFNHNDNQTM 180
QY 180 RLVCMLYTPLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVETKTPFVSNLPL 239
DB 181 RLVCMLYTPLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVETKTPFVSNLPL 238
QY 240 NTLNSRVPSLIKSMVSRDHQVQFONGRVTLTGOLGTTPTTSASOLCKIRGSVFHAN 299
DB 239 SLSNSRAPLPISMSGISPDNVQSQFQNGRCVTLGRLVGTTPVSLSHVAKIRGT---S 294
QY 300 GGGYNLTLDGSPYHAFESPAPIGPDLGECDDHMEASPTTFQNTGVDVVKQBSA 359
DB 295 NGVINLTLDGTPFPFPEGPAIGPDLGGCDWHIN---MTQFGHSSQTVQDVTTPDT 351
QY 360 FAPHLGTQADGLSDSVSNNTAKLGVSPVSDGHRGVDVDPVTPRGSTLTLEAAQLAP 419
DB 352 FVPHLSIQANGIG---SGNYGVLSWISPPSHPSGSQVDLWKIPNNYSSITEATHLAP 407
QY 420 PIYPGGEIIVFMSDFPIAHGTNGLSVPCTIPQEFVTHFVNEQAPTRGEAALLHYLDP 479
DB 408 SVYPPGGEIVLVFMSQMP---GPGAYNLECLLPQEVISHLASEQAPTVEAALLHYVDP 464
QY 480 DTHRNLEFKLYPEGFMTCPVNSGSGTPQPLPINGVVFVSWVSRYQLKPVGTAGPA-C 538
DB 465 DTGNLGEFKAYPDGFLTCVPNGASSGPQQLPINGVVFVSWVSRYQLKPVGTASSARG 524
QY 539 RLGIIR 544
DB 525 RLGIIR 530

RESULT 2
US-10-670-695-4
; Sequence 4, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; PRIOR FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 671
; TYPE: PRP
; ORGANISM: Feline calicivirus
US-10-670-695-2

Query Match      9.5%; Score 276; DB 12; Length 671;
Best Local Similarity 26.4%; Pred. No. 2.6e-17;
Matches 102; Conservative 54; Mismatches 128; Indels 102; Gaps 17;

QY 13 DGTSGAGQLVPEANT-AEPISEMPVAGATAAATAGQVNMIDPWIMNMYVQAPQGEFTI- 70
DB 127 DGSITA---PEQGTWVGVIAPESQAQMSAADMATGKSVSEW-----EAFPSFH 173
QY 71 -----SPNPTPGDILDLQGLPHNPFSLHQAQMYNGVGNMKVYKVLNAGNFTAAGKII 125
DB 174 TSVNWTSETQKILFKQSLGFLNLPYLEHLAKLVANSGSIEVRFSISGSGVFGGLAA 233

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 547
;   TYPE: PRP
;   ORGANISM: Feline calicivirus
US-10-670-695-4

Query Match      9.5%; Score 276; DB 12; Length 547;
Best Local Similarity 26.4%; Pred. No. 1.9e-17;
Matches 102; Conservative 54; Mismatches 128; Indels 102; Gaps 17;

QY 13 DGTSGAGQLVPEANT-AEPISEMPVAGATAAATAGQVNMIDPWIMNMYVQAPQGEFTI- 70
DB 3 DGSITA---PEQGTWVGVIAPESQAQMSAADMATGKSVSEW-----EAFPSFH 49
QY 71 -----SPNPTPGDILDLQGLPHNPFSLHQAQMYNGVGNMKVYKVLNAGNFTAAGKII 125
DB 50 TSVNWTSETQKILFKQSLGFLNLPYLEHLAKLVANSGSIEVRFSISGSGVFGGLAA 109
QY 126 SCIPGCF-AAQNISIAQATMFPHVIADVRVLEPIEVLDPVRLVFNHNDNAPTMRLVCM 184
DB 110 IVVPGVDVQVSTMLQ---YPHVLFDAQVPEFVFCPLDRLKSTLYHLMSDITDTSLVIM 166
QY 185 LY-----TPLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVETKTPFVSNLPLN 240
DB 167 VYNDLINPYANDANSSGC---IVT--VETKPGDFKFKHLLKPPG-----SMLTHG 211
QY 241 TLSNSRVPSLIKSMVSRDHQVQFONGRVTLTGOLGTTPTTSASOLCKIRGSVFHAN 300
DB 212 SIFSDILPKTSSLIWGNVMSDITDP-----VIRPFVFOANR 248
QY 301 GNGYN-----LTLDGSPYHA-----PESPA-PIGFPDL---GEC--- 331
DB 249 HFDNFOETAGWSPRFRPISVTITEQNGAKLIGVATDVIIVGIPDGWPTTIPGELIPA 308
QY 332 -DWHM-----EASPTTFQNTGVDVVK 350
DB 309 GDYAITNGTNDITATGVDYTDIHK 334

RESULT 3
US-10-670-695-2
; Sequence 2, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; PRIOR FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 671
; TYPE: PRP
; ORGANISM: Feline calicivirus
US-10-670-695-2

Query Match      9.5%; Score 276; DB 12; Length 671;
Best Local Similarity 26.4%; Pred. No. 2.6e-17;
Matches 102; Conservative 54; Mismatches 128; Indels 102; Gaps 17;

QY 13 DGTSGAGQLVPEANT-AEPISEMPVAGATAAATAGQVNMIDPWIMNMYVQAPQGEFTI- 70
DB 127 DGSITA---PEQGTWVGVIAPESQAQMSAADMATGKSVSEW-----EAFPSFH 173
QY 71 -----SPNPTPGDILDLQGLPHNPFSLHQAQMYNGVGNMKVYKVLNAGNFTAAGKII 125
DB 174 TSVNWTSETQKILFKQSLGFLNLPYLEHLAKLVANSGSIEVRFSISGSGVFGGLAA 233

```


QY 126 SCIPGF-AAQNISIAQATMPHHVIAVVRVLEPIEVPLEVVRNVLFNHNDNAPTMRVLCM 184
DB 234 IWPFGVDPVQSTMLQ---YPHVLFDAQVEPVIFCLPDLRSTLYHLMSDITDTTSLVIM 290
QY 185 LY-----TPLRSGSSGCTDPFVIAGRVLTCPSPDFSLFVLPVPPNVEQTKFSPVNLPLN 240
DB 291 VYNLDLINPYANDANSNGC---IVT---VETKPGDPFKFHLKPPG-----SMLTHG 335
QY 241 TLSNRVPSLIKMMVSRDHGMQVQFNGRVLTLDQLOQTTPTSASQLCKIRGSVFHANG 300
DB 336 SIPSLIIPKTSLSLWIGNYWSIDTF-----VIRPFVQANR 372
QY 301 GNGYN-----LTLDGSPYHA-----FESPA-PIGPPDL---GEC--- 331
DB 373 HFDFNQETAGWSTPRFRPITVTEQNGAKLIGVATDVIVGIPDGWPDPTTIPGELIPA 432
QY 332 -DWM-----EASPTQFNTGDVTK 350
DB 433 GDYAITNGTNDITATGYDTADIK 458

RESULT 4

US-10-209-507-2
; Sequence 2, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:

; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus

US-10-209-507-2

Query Match 9.4%; Score 272.5; DB 14; Length 669;
Best Local Similarity 25.9%; Pred. No. 5.6e-17;
Matches 106; Conservative 54; Mismatches 134; Indels 115; Gaps 19;

QY 13 DGTSGAGQLVPEANT-AEPISEMPVAGATAAATAGQVNMIDPIMNMYVQAPQGEFTI- 70
DB 126 DGDSSI--TTPEQTLVGGVIAEPNAQMSAVADVATGKSDSEW-----ESFFSFH 174
QY 71 -----SPNNTPGDILFDLQGLPHLPFLSHLAQMYNGVGNMKVLLAGNAFTAGKIII 125
DB 175 TSUNWSTSETQKILFKOSLGLPLNPLYLHLSKLYVANGSGVDVRFSSGSGVGGKLA 234
QY 126 SCIPGF-AAQNISIAQATMPHHVIAVVRVLEPIEVPLEVVRNVLFNHNDNAPTMRVLCM 184
DB 235 IWPFGVDPVQSTMLQ---YPHVLFDAQVEPVIFSLPDLRSTLYHLMSDITDTTSLVIM 291
QY 185 LY-----TPLRSGSSGCTDPFVIAGRVLTCPSPDFSLFVLPVPPNVEQTKFSPVNLPLN 240
DB 292 VYNLDLINPYANDANSNGC---IVT---VETKPGDPFKFHLKPPG-----S 331
QY 241 TLSNRVPS--LIKSMV---SRDHGMQVQFNGRVLTLDQLOQTTPTSASQLCKIRGSV 295
DB 332 MLTHGSIFSDLPKSSSLWIGNRWTSIDTF-----VIRPFV 368
QY 296 FHANGNGYN-----LTLDGSPYHA-----FESPA-PIGFPD--LGE 330

DB 369 FOANRHDFNOETAGWSTPRFRPITITISENSGSKLGTGVATDVIVGIPDGWPDPTTIGE 428
QY 331 CDWHEASPTTQFNTGDVTKINVKQBSAFAPHLGTIQADGLSDVSVNVT 379
DB 429 -----ELTPAGDYSI-----TNGSGNDIATANAYDSADVITNT 461

RESULT 5

US-10-209-507-4
; Sequence 4, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:

; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus

US-10-209-507-4

Query Match 9.4%; Score 272; DB 14; Length 668;
Best Local Similarity 26.7%; Pred. No. 6.3e-17;
Matches 107; Conservative 46; Mismatches 134; Indels 114; Gaps 20;

QY 23 PEANT-----AEP-ISMPEVAGATAAATAGQVNMIDPIMNMYVQAPQGEFTI----- 70
DB 133 PEQTLVGGVIAEPNAQMSAVADVATGKSDSE-----W-----EAFSPHSTSVN 177
QY 71 -SPNNTPGDILFDLQGLPHLPFLSHLAQMYNGVGNMKVLLAGNAFTAGKIIISCIP 129
DB 178 WSTSETQKILFKOSLGLPLNPLYLHLSKLYVANGSGIEVRFSSGSGVGGKLAIVVP 237
QY 130 PGF-AAQNISIAQATMPHHVIAVVRVLEPIEVPLEVVRNVLFNHNDNAPTMRVLCMLY-- 186
DB 238 PGIDPVQSTMLQ---YPHVLFDAQVEPVIFITPDLRNSLYHLMSDITDTTSLVIMIYND 294
QY 187 --TPLRSGSSGCTDPFVIAGRVLTCPSPDFSLFVLPVPPNVEQTKFSPVNLPLNTLSN 244
DB 295 LINPYANDANSNGC---IVT---VETKPGDPFKFHLKPPG-----SMLTH 334
QY 245 SRVPS-LI---KSMVSRDH--GMVQFQNGRVLTLDQLOQTTPTSASQLCKIRGSVFHAN 299
DB 335 GSIPSDLPKSSSLWIGNRWTSIDTF-----VIKPFVFOAN 371
QY 300 GNGYN-----LTLDGSPYH---AFESPAPIGFDLGECDWHMEASP 339
DB 372 RHDFNOETAGWSTPRFRPITITVSEKSGSKLGTGVATDSIVP-GIPD-----GWPDTTIP 426
QY 340 TTQFNTGDVTKINVKQBSAFAPHLGTIQADGLSDVSVNVT 380
DB 427 EKLTPAGDYAITNGGNDITTA-----ADYDGASIIKNTN 462

RESULT 6

US-10-282-122A-67793
; Sequence 67793, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu

```

; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 67793
; LENGTH: 6310
; TYPE: PRT
; ORGANISM: Pseudomonas putida
; US-10-282-122A-67793

Query Match      4.2%; Score 122.5; DB 12; Length 6310;
Best Local Similarity 19.5%; Pred. No. 0.67;
Matches 110; Conservative 70; Mismatches 221; Indels 163; Gaps 22;

QY      8  APTNMDGTSAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNMYVQAQGE 67
D      4016  AAGNLSNLSAATVLSGGGEGAGATVVRDASGAILATGTVNQSQFOITLPSAQVTGSLQ 4075
QY      68  FTIS-----PNNTEGIDILDLQLGPHLNPFLSHLAQMYNG-----WVG 105
D      4076  VTLSDAAGNVSGPASLATPDHTPPAAI-----SNPVLSDGROGLSGSGEAGATVQVR 4127
QY      106  NMKVVLGNAGTAFTAGKIIISCIPPGFAAQNISIAQATMPPHYADVRLIEVPLEDV 165
D      4128  NAAGALLGTATVGTGDRFTVTFDTPQANGQVIGVTOMDAASNTSPAINVTPDLTPAPL 4187
QY      166  RNVLFNHNDNAPTMRVLCMLYTPLRASGSS-----SGTDPFVI----- 203
D      4188  TNNVLNNG-----LTLTGEGEAGATVTVHGPDTIIGTGLVAANGSFTLILNSAQ 4238
QY      204  -----AGR-----VLTCSPDPSFLFLVPPNVEQTKP--FSVNLPLNTLSNS 245
D      4239  LNAQLLSVTQTDAGNNISTAVATAPDFT-----PP-----TAPTALALSGTGLQTGNA 4288
QY      246  RVPSLSKMMVSRD-HGQMVFQNG--RVTL-----DQLOGTTPTTASQLCKIRGVSFVH 297
D      4289  EAGSTVTVRDSAGNVLTGAVAGNGTFTQVLNSAQTHGQILQVLTATDAA-----GNVSP 4342
QY      298  ANGGNGYNLTDLGDSPHYAFESPAPIGFPDLEGEC-----DWHMEASPTTQNTGDVI 349

```

```

Db      4343  A-----APYTAADTTTPAAVANLAVSANGATLTGDCGAGATVTVRAPDGTVL 4389
QY      350  KQINVKQESAF-----PHLGTTOADGLSDVSNTNMIAKLGHVSVSDGHRGDV 399
D      4390  GNATVAADGHFVSLSPAAITGESLSVVQDAQAQNVSPAQNTAP-GALAAATP----- 4442
QY      400  DPWVTPRYGSTLTTEAAQLAPPIYPPGFGEAIVFFMSDFPIAHGTNGL---SVPCTIPOEF 456
D      4443  DNLILAADGLSVSGTAEGSTIK-----VYGPNGVLLGSSPVTNDGTF 4485
QY      457  VTFHVNQAPTGRGEAALLHYLDDP 480
D      4486  TVNLGSAQA--NGEVLQVSATGPD 4507

RESULT 7
US-10-369-493-2294
; Sequence 2294, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2294
; LENGTH: 913
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
; NAME/KEY: unsure
; LOCATION: (1)..(913)
; OTHER INFORMATION: unsure at all Xaa locations
; US-10-369-493-2294

Query Match      4.2%; Score 122; DB 15; Length 913;
Best Local Similarity 19.5%; Pred. No. 0.037;
Matches 116; Conservative 90; Mismatches 226; Indels 164; Gaps 30;

QY      2  MWASKDAPTNNMDGTSAGQLVPEANTAEPISEME-----PVAGAATAAA-----TAG 47
D      238  IMVSTVTPAVGSTTSTVDPDVVFXAGNVEVYPTEDPDIMVSVTPAVGSTTSTVDPDVVPXAG 297
QY      48  --QVNMIDPWIMNMYVQAQGEFTI-SPNNTPGDILFDL-OLGPHLNPFLSHLAQMYNGW 103
D      298  IVEVNGIDPEITVSVTPAVGSTTMTAPDEVSKXAVIVDIKETDPEIT--VSVTPAVGST 355
QY      104  VGNMKVKVLLAGNAP---TACKIIISICPPGFAAQNISIAQATMPPHYI-----ADVRVL 155
D      356  TSTVTRVVPVLAGNVEVYPTEDPDIMVSVTPAVGSTTSTV-----PEVPLAGNVEVVKVL 409
QY      156  BP---IEVPLEDVRN-----VLFHNDNAPTMRVLCMLYTPLRASGSSGSDTDPFVI-- 203
D      410  EPTMTVSVIVPAVGSTTSTVPEVPLAGNVEVNALEPDTMVSIVPAVGSTTSTVPEVVP 469
QY      204  -AGRYLTCSPDPSFLFLVPPNVEQTKPFSVFN-LPNTLSNRSRVPSLKSMVMVSRDHG 261
D      470  LAGNVEVYPTEDPDIMVSVTPAVGSTTSTV-----TVPEVPL--AGNVEVYPTEDPDIMVSV 521
QY      262  QMVQONGREVTLGDLGQLGTTPTASQLCKIRGVSFVHANGNGYNLTDLGDSPHYAFESPA 321
D      522  -----VTVPA--VGSSTSTVPEVPLAG-----NVEVYPTEDPDIMVSV 555
QY      322  PIGFDPDLGDCDWHMEASPTTQNTGDVIKINVKQESAFAPHLGTIOADGLSDVSNTNM 381

```

Db 556 SVIVPAVG-----STNSTVPEV-----VPLAGNVEVNALEPDTMVSVI 593
Qy 382 IAKLG-----WYSPVSDGHRGDVD-----PWVIRYGSTLTERRAQLAP----- 419
Db 594 VPAVGSTTSTVPEVPLA-----GNVEVKVLEPETMVSIVPAVGSTTSTVPEVPLAGNV 649
Qy 420 PIYP-----PGGEAIVFFMSDFPIAHGTNGLSVPCTIPQEFV-----T 458
Db 650 EYVTEPETMVSIVPAVGSTTSTVPEVPLA-----GNVEVKVLEPETMVSIVPAVGSTT 706
Qy 459 HFVNEQAPTRGEAALLHVLDDTHNLGELFKLYPEGFMTCVPNSGGTGCPQTLPLING 514
Db 707 STVPEVPLAGNVE-VNALEPDTMVS-----IVPAVGSTTSTVPEVPLAG 752

RESULT 8

US-10-282-122A-77663
; Sequence 77663, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Malone, Cheryl
; APPLICANT: Olsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-09-09
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2001-02-16
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77663
; LENGTH: 3263
; TYPE: PRT
; ORGANISM: Vibrio cholerae

US-10-282-122A-77663
Query Match 4.2%; Score 122; DB 12; Length 3263;
Best Local Similarity 20.0%; Pred. No. 0.27;
Matches 120; Conservative 71; Mismatches 197; Indels 212; Gaps 30;

Qy 25 ANTAEPISMEBVAGA-----ATAATA--GQVNMIDPWI-----MNNYVOAPQ 65
Db 1746 AQVVTGIVQVQVAVASSTANIDIKATVITVENDGHRVQVEETIRILVEPKIDVTENYHNAV 1805

Qy 66 G-----EFTISPNTTDFILFDLQLGPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFT 119
Db 1806 GNEDDRIHVTVWPQNTPGNI-----QNPD-----AQEY----- 1833
Qy 120 AGKIIISCIPG--FAAQNISIAQATMPPHVADRVLEPIEVPLE-DVRNVLP----- 170
Db 1834 FSRVEISGFDPGSRVFNNEVT-----LINGVLVLEPAAGQSDLDLFSNQVSAAGYIQ 1886
Qy 171 ----HNDNAPTMLVCMW----YTPLRASGSSSGTDPFVIAG----RVLTCPSPDFSL 218
Db 1887 VIPHNSSTDTLTSTAITVKEQDEHYVDAGNPGOGIAEEVTHGSGIKVKNPIAEFDGQL 1946
Qy 219 FLVPEVNEQTKPFSPNPLPLNTLSNRVPSLISKMMVSRDHGMVQFQNGRVTLDGQLQ 278
Db 1947 VENAGSVTQ-----TVQADANGKIDFTINNVSQGGQGANVIRFDNLDSNTAGSYQ 1996
Qy 279 -----GTTPTSASQLCKIRGSVPHANGNG-----YNLTGDSPPYHAF 317
Db 1997 SDELVDQLVVSFGNVPOEVLNQLLITGAI---NNGDGTWTITNBADFSIKAPNGLVYSSN 2053
Qy 318 ESPAPIGFPDLGECDWMEASPTTQ-FNTGD-----VIQINVKQESAFAPHLGTIQAD 370
Db 2054 NDPKNGFNDI-----KITITAKVYDQGEDSSEVKITKOVSTELTSLFP----- 2097
Qy 371 GLSDVSVNTNMIAKLGMVSPVSDGHRGDVDPWVIRYGSTLTAAQLAPP--IYPPGFGE 428
Db 2098 --TEVTGNNVAAQLNWW-----GDADDLVIGKEDNTVNLGQIQDKLMVNATGPD 2147
Qy 429 -----AIVFFMSDPPIAHGTNGLSVPCITIQEFVTHFVNEQAPTRGEALLHYLDPDTHR 483
Db 2148 VADELSIVINASDLFAGASIGG-----QDF--NFVDGHVYFKGT-----LNPDSGI 2191
Qy 484 N-----LGEFKLYPEGFMTC-----VPNSSSGTGP 507
Db 2192 SGLEGLVLPDRPAGDFKL-PITFVTITDTSQSGDEKTLTAQVPAISPVADVPSSGDP 2250

RESULT 9

US-10-282-122A-48055
; Sequence 48055, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Olsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-09-09
; PRIOR FILING DATE: 2000-09-09
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27

US-10-282-122A-48055
Query Match 4.2%; Score 122; DB 12; Length 3263;
Best Local Similarity 20.0%; Pred. No. 0.27;
Matches 120; Conservative 71; Mismatches 197; Indels 212; Gaps 30;

Qy 25 ANTAEPISMEBVAGA-----ATAATA--GQVNMIDPWI-----MNNYVOAPQ 65
Db 1746 AQVVTGIVQVQVAVASSTANIDIKATVITVENDGHRVQVEETIRILVEPKIDVTENYHNAV 1805

Wed Jun 2 09:13:23 2004

```

; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 48055
; LENGTH: 1765
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-282-122A-48055

Query Match      4.0%; Score 114.5; DB 12; Length 1765;
Best Local Similarity 20.9%; Pred. No. 0.55;
Matches 102; Conservative 58; Mismatches 198; Indels 129; Gaps 22;

QY 2 WKASKDAPTN-----MDGTSGAGQLVPEANTAEPISE-----PVAGATAAAT-----AG 47
Db 966 MIAPTEPTSSVTPMTGAPG-----VVSPTVTTEADGVLPAASVAVTPTTVPAG 1017
QY 48 QVNMIDPWTMNNYVQAPOGEFTISPNNTGDLFDLQGLPHLNPPLSHLAQWYNGWYGNM 107
Db 1018 S-----GVEGVVHAPLGSVAVTASGVP-----SPFVSMLTAVAPASAVPDS 1058
QY 108 KVKVLLAGNAFTAGK--IIISCIPGFAAQNISIAQATMFPFHVIADVRVLEPIEVPLEDV 165
Db 1059 DVPSV-----AFTVGNAGAVVSIVSKAVLGTTLTLPAS-----VAVTVRLCAPSPKPVGV 1110
QY 166 RNVLFHNDNAPTMRILVCMLYT-----PLRAGSSSGTDPFVIAGRVLTCTCPSPDFS 216
Db 1111 -NVQFPGSVAVVPNNVPSYTLTLPASAVPLNVGVASSVLPPEMIA----- 1157
QY 217 FLFLVPNVQKTPSPVNPPLNTLSNRVPSLIKSMVSRDHGQMVQFQNGRVTLGQ 276
Db 1158 -----PTTEPTSSVTPMTGAPG-----AVSPVTTEADGVP-----LPAA 1194
QY 277 LOGTTPTSASQCKIRGVSFHANGNGYNLTLDGSPYHAFESPAPI-GFPDLGCDWHM 335
Db 1195 SVAVTPTVPAGSGVEGVVHAPLGSVAVTASGVPSPFVSMLTAVAPASAVPD----- 1246
QY 336 EASPTTQNTGDIKQINVKQESAFAPHLGTIOADGLSDVSNTNMIK-----LGWVSP 390
Db 1247 NDVPSVAVTGVLAGAVSVIVRSKAV---LGTLLTPAVS-VAVTVRLCAPSPSNAVGVNQV 1302
QY 391 VSDGHRGDVDPWIPRYGSTLTAAQ-----LAPPIYPPGGEAIVFFMSDFPIAHGTN 444
Db 1303 FPDGSVAVVPNNVPSYTLTLPASAVPLNVGVASSVLP-----EATA-----PTTERAS 1353
QY 445 GLSVPECT 451
Db 1354 SVTVPTMT 1360

```

```

RESULT 10
US-10-156-761-9262
; Sequence 9262, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089

```

```

; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9262
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9262

Query Match      3.9%; Score 113; DB 14; Length 724;
Best Local Similarity 20.5%; Pred. No. 0.19;
Matches 130; Conservative 48; Mismatches 220; Indels 236; Gaps 28;

QY 14 GTSAGAGQLVPEANTAEPISEMEPVAGATAAATAAGVNMID-----PWIMNNYVQ 62
Db 160 GDKIAVEAVSTQTVVDVADFEOVAAAAT--KPAGAVSVTDKADPSGGGDSGTQAFRNAIA 217
QY 63 APOGEFTISPNNTGPDILFDLQ-----GPHLNP 91
Db 218 AAGGVVWTP---PGDYRLTSSLNGVQNVTLQAGSMHVVHTSRFIDQSSSSCGVHLKD 274
QY 92 FL-----SHLAQWYNG-----WVGNMVKVLLAGNAFTAGKIIISCIPP 130
Db 275 FAVIGEVTERVDSNPDNFVNGSLGPGSSVSGMWLOHLKVGLWLTGN-----ND 322
QY 131 GFAAQNISIAQATMFPFHVIAD-----VRVLEPIEVPLEDVNRNVLFHNDNAPTMR 180
Db 323 NLVVENNRILDT-----ADGLNLNGNARGVRVR-----NNFLRNGDDALAM- 365
QY 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCTCPSPDFSFLVPPNVQKTPSPVNPPLN 240
Db 366 --NSLYSPDNTSSFSNT-----ISQPNLANG 390
QY 241 TLNSNRVPSLIKSMVSRDH--GQVQFONGRVTLG--GQLOQTTPTSASQCKIRGVSF 296
Db 391 IAYGGTDLAVKKNLISDNTVALGSGIAISNOKE-LDPFSPLAGTITVSGNTL--VRAGAM 447
QY 297 HANGNGYNI TELDG-----SPHAFESPAPIGFPDLGCDWHMEAS 338
Db 448 NPNQHPMGALRVDSYDSEANVSITGTTTSPYSAFEVSGGGH---GYATKNTVVD 504
QY 339 PTTQNTGDIKQINVKQESAFAPHLGTIOADGLSDVSVN-----TNMIK 385
Db 505 GATVRNTGTW---VOAESQGAAPSNVQRTGGAAGIYNCYPYAGSGTFTLTDCGGNA 560
QY 386 GWVSPVSDGHRGDVDPWIPRYGSTLTAE-----AAQLAPPIYPPG---FGEAIV 431
Db 561 GWSST-----WGDCSAMPQCGQNPDPDPTRNLAKGRPATATGSDVYTFPKAVDGDAST 615
QY 432 FMSDFPIAHGTNGLSVPCPTIPQEFVTHFVNEQAPTRGEAALLHYLDDPDTHRLNGEPKLY 491
Db 616 YWES-----TNN-----APPOAWTVDLGSSQAVRR---LVKLKLPATAWQARTQTL 659
QY 492 PEG-----FMTCV-----PNSSGTGPOTLP 511
Db 660 VQGSTGSAISTVVASQGYRFPDPATGNTATVTLTP 693

```

```

RESULT 11
US-10-128-714-3298
; Sequence 3298, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroehkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999

```

```
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3298
; LENGTH: 1555
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-3298

Query Match      3.9%; Score 113; DB 14; Length 1555;
Best Local Similarity 19.2%; Pred. No. 0.63;
Matches 104; Conservative 61; Mismatches 177; Indels 200; Gaps 28;

Qy  2 MMASKDAPNTMDGTSGAGQLVPEANTAPISMVPVAGAAATAAGQVNMIDPIMNNYV 61
Db  525 LKASTVSPAPL-GTSSAPTAPQKNFFBELPLPPRPKSRPASSG-----RYTPNAPV 576

Qy  62 QAPQGEFTISP-----NNTPGDILFDLQGLHPLNPFSLHQAQMYNGVGNMKVKVLLAGN 116
Db  577 SAPLSQSIPPANQYSNVPG-----APQSN-----IG----- 604

Qy  117 AFTAGKIIICITPPGFAAQNISIAQATMPFHVIADRVLEIEVPLEVDVNVLFHNDNA 176
Db  605 -----PPD-----PPLOQPERLDPS-----NLLAPNVSA 631

Qy  177 PTMLVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVVEQTKPFSPVN 236
Db  632 PAVPSTASRYSP-RPPGVQAGVKP-----PPSPRYS-----PAPPQSTN--AVAA 673

Qy  237 LPLNTLSNSRVPSLIKSMVSRDHQMVQFQNGRVTLDGQLQGTTPTSASOLCKIRGSVF 296
Db  674 APRNRYASQ--PASISG-----QGAALQFO-----PRTSSPLA-YHEKIH 710

Qy  297 HANGGNGYLTDLGSPYHAFESAPIGFDPDLGECDMHMEASPTTQFNTG----DVIKQI 352
Db  711 YEDQGSERQLOST-----ASPPPLNHS-----HPSEQPVSSSENKGSQVDVLENV 758

Qy  353 -----NVKQESAPAPHLGTIQADGLSDVSVNTNMIAKLGWVSPVSDG----- 394
Db  759 PPLSTRQSPKPNYPAP-----SAYTNEFANR--VAPVSTGPPGIAGMTGVLN 803

Qy  395 HRGVDVDPWIVPRYGSITLTAQAQLAPPYPPGFGGAI VFFMDSDFPIAGHTNGLSVPTPIQ 454
Db  804 SSTEESPVPPRRSQTSQSPQLSPRLSVPS-----LDPFORPASVHGSTSP----- 850

Qy  455 EFVTHFVNEQAP-----TRGEAALLHYLDPDTHRNLGFEK-----LYPSEGF----MTC 498
Db  851 ---TRTVNYPAPVPVPTHNRAPSVQLEFIPTDCCQLDLSLRWKGAPIFKFGFGNAVISC 907

Qy  499 VP 500
Db  908 FP 909
```

RESULT 12

```
US-10-128-714-8298
; Sequence 8298, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
```

```
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 8298
; LENGTH: 1832
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8298

Query Match      3.9%; Score 113; DB 14; Length 1832;
Best Local Similarity 19.2%; Pred. No. 0.82;
Matches 104; Conservative 61; Mismatches 177; Indels 200; Gaps 28;

Qy  2 MMASKDAPNTMDGTSGAGQLVPEANTAPISMVPVAGAAATAAGQVNMIDPIMNNYV 61
Db  525 LKASTVSPAPL-GTSSAPTAPQKNFFBELPLPPRPKSRPASSG-----RYTPNAPV 576

Qy  62 QAPQGEFTISP-----NNTPGDILFDLQGLHPLNPFSLHQAQMYNGVGNMKVKVLLAGN 116
Db  577 SAPLSQSIPPANQYSNVPG-----APQSN-----IG----- 604

Qy  117 AFTAGKIIICITPPGFAAQNISIAQATMPFHVIADRVLEIEVPLEVDVNVLFHNDNA 176
Db  605 -----PPD-----PPLOQPERLDPS-----NLLAPNVSA 631

Qy  177 PTMLVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVVEQTKPFSPVN 236
Db  632 PAVPSTASRYSP-RPPGVQAGVKP-----PPSPRYS-----PAPPQSTN--AVAA 673

Qy  237 LPLNTLSNSRVPSLIKSMVSRDHQMVQFQNGRVTLDGQLQGTTPTSASOLCKIRGSVF 296
Db  674 APRNRYASQ--PASISG-----QGAALQFO-----PRTSSPLA-YHEKIH 710

Qy  297 HANGGNGYLTDLGSPYHAFESAPIGFDPDLGECDMHMEASPTTQFNTG----DVIKQI 352
Db  711 YEDQGSERQLOST-----ASPPPLNHS-----HPSEQPVSSSENKGSQVDVLENV 758

Qy  353 -----NVKQESAPAPHLGTIQADGLSDVSVNTNMIAKLGWVSPVSDG----- 394
Db  759 PPLSTRQSPKPNYPAP-----SAYTNEFANR--VAPVSTGPPGIAGMTGVLN 803

Qy  395 HRGVDVDPWIVPRYGSITLTAQAQLAPPYPPGFGGAI VFFMDSDFPIAGHTNGLSVPTPIQ 454
Db  804 SSTEESPVPPRRSQTSQSPQLSPRLSVPS-----LDPFORPASVHGSTSP----- 850

Qy  455 EFVTHFVNEQAP-----TRGEAALLHYLDPDTHRNLGFEK-----LYPSEGF----MTC 498
Db  851 ---TRTVNYPAPVPVPTHNRAPSVQLEFIPTDCCQLDLSLRWKGAPIFKFGFGNAVISC 907

Qy  499 VP 500
Db  908 FP 909
```

RESULT 13

```
US-10-093-463-82
```

; Sequence 82, Application US/10093463
; Publication No. US20030208039A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shenoy, Sureesh
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Gusev, Vladimir
; APPLICANT: Pochart, Pascal
; APPLICANT: Zhong, Mei
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter
; APPLICANT: Smithson, Glenda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gerlach, Valerie
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Tchernev, Velizar
; APPLICANT: Gangolli, Esha
; APPLICANT: Vernet, Corine
; APPLICANT: Pena, Carol
; APPLICANT: Burgess, Catherine
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gorman, Linda
; APPLICANT: Spaderna, Steven
; APPLICANT: Voss, Edward
; APPLICANT: Malyankar, Uriel
; APPLICANT: Anderson, David
; APPLICANT: Patturajan, Meera
; APPLICANT: Miller, Charles
; APPLICANT: Taupier, Raymond J. Jr.
; TITLE OF INVENTION: No. US20030208039A1e1 Antibodies that Bind to Antigenic Polypeptide
; TITLE OF INVENTION: Encoding The Antigens, and Methods of Use.
; FILE REFERENCE: 21402-230A (Cura 590AT)
; CURRENT APPLICATION NUMBER: US/10/093,463
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/283,675
; PRIOR FILING DATE: 2001-04-14
; PRIOR APPLICATION NUMBER: 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,101
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/287,424
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/299,027
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/309,198
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/281,194
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/330,380
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/288,342
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/275,578

; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 82
; LENGTH: 2914
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-093-463-82

Query Match 3.9%; Score 112.5; DB 15; Length 2914;
Best Local Similarity 21.3%; Pred. No. 1.9; 234; Indels 183; Gaps 33;
Matches 132; Conservative 71; Mismatches

QY 5 SKDAPNTMDGTSGAGQLVPEANTAEPISEMPVAGAAATA---AATAGQVNMIDPWIMNYV 61
DB 1473 SQGLTPSDEEPQLSQESFRTTTRP-ALTPAAPLTTALNPEVTATEPVPVSPGQTOTL 1531
QY 62 QAPQGEFTTSPNNTPGDILFDLQGLHNPFLSHLAQMYNGWGNMKVKVLLAGNAFTAG 121
DB 1532 QQFL-ELTAS-----QLPAGFTESP-----ASKGVTA 1558
QY 122 KIIISCIPGFAAQNISIAQATMEPHVIADVRLPEIEVPLEDVR-NVLFHNNDN----- 175
DB 1559 LLAIPHTP---ESSSLFVALQITPTQMVSGA-----METTRVTIVIFAGSPNITVSS 1606
QY 176 ---APTMRILVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVQKTKP 231
DB 1607 RSPPAFRFPLMTKAVT-VRGHGS-----LPVRT-TPQPFSLT-----ASPSSRRVASP 1652
QY 232 FSVNPLPLNTLSNSRVPSLIKSMVSR-----DHQMVQFQNGRVTLDDQLQGTTPTS- 284
DB 1653 GAISRSPTSSGSHKAVLTTPAVTKVISRTGVPOPTQOASASSPSTPLTVAGTAAEQVPVSP 1712
QY 285 -ASQLCKIRGSVFHANGNGYNLTLDGSPYHAFESPAPIGFPDGLGECDMHWEASPTTQF 343
DB 1713 LATRSLEIVLSTKGEAGHSQPM---GSP---ASFPQHPPLP-----SAPPRPAQ 1755
QY 344 NTGDVIKQINVKQESAPAPHLGTIQADG-----LSDVSVNTNM-IAKL 385
DB 1756 HTTMATRSALPPETPAAASLST-ATDGLAATFMSLESTRPSQLLSGLPDTLSLAKV 1814
QY 386 GWSPV-SDGHRGDVDWVPIRYGSTL-----TEAA--QLAPP--IYPPGFGEA 429
DB 1815 GTSAPVATPGPKASVITTPLOQOATTLPAQTLSPVLPTTAAATQAHPPTHIAPPAAGTA 1874
QY 430 IVFFMSD-----FPTAHGTNGLSVCTTIPQEFVTHFVNEQAPTGERALLH----- 475
DB 1875 PGLLLGATLPTSGVLPAEGLT--ASMVSVVPRK-----STTKVAILSKQVSLPT 1922
QY 476 -----YLDPDTHRNLGEPKLYPEGFM-----CVFNS---SGTGPQTLPINGVFV 517
DB 1923 SMYGAEGGPTELTPTATSHPLTPLVAEPGAQAQATLVPVTSVLSRVSAARTAPQDSMLV 1982
QY 518 FVSWVSRYQLKPVGT-AGP 536
DB 1983 LLPQLAEAH-----GTSAGP 1997

RESULT 14

US-10-282-122A-46817
; Sequence 46817, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert

```
Query Match      3.9%; Score 112; DB 12; Length 3930;
Best Local Similarity 18.8%; Pred. No. 3.4;
Matches 135; Conservative 86; Mismatches 264; Indels 232; Gaps 34;

Qy   5 SKDAP-----TNMDGT-----SGAGQLVPEANTAEPIISMEPVAGAATAA-----43
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db   746 SAUSPTFVDNPDGTTFTPINTFLINGVLQNADPNVGVPPLPSIFANGSLTVSYQVTVTSL 805

Qy   44 -----ATAGQVNMI-----DPWMN-----NYVQAPQGEBTI-----70
    ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db   806 PTQNPTINSSSTQVSFILNPGDPPPTIETLSLNTVSTQILNANVVIVKQVDLTIAADVGPDI 865

Qy   71 -----SPNNTPGD---ILFD-IQLGPLHLP---FLSHLAQMNGYGVGNMKVKVLLAGNA 117
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 15

359 GFVPHLGSIQFDEVFNHPTG-DYIGTIEWISQSPSTPPTDINLWEIPDYGSSISQAANLA 417
 419 PPIVPPGGEAIVFMSDFPPTAHGTNGLS----VPCITPOEFVTHFVNEQAPTRGEAALL 474
 418 PVVPPGGEALVYFVSAPP---GNNKRSAPNDVPCLLPQBYIYTHFVSEQAPTWGDAALL 474
 475 HYLDPTHRLNLFGEFKLYPEGMTCVPNSSGTGPTLPINGVFFVSVWSRFPYQKPVGTA 534
 475 HYVDPDTHRLNLFGEFKLYPGGYLTCVGVGAGPQQLPLNGVFLFVSVWSRFPYQKPVGTA 534
 535 GBA-CRLGIRR 544
 535 STARGRLGVR 545

RESULT 2
 B37471
 capsid protein - Norwalk virus
 C:Species: Norwalk virus
 C:Date: 24-Feb-1994 #sequence_revision 25-Apr-1997 #text_change 28-Jul-2000
 C:Accession: B37471
 R:Yang, X.; Wang, M.; Wang, K.; Estes, M.K.
 Virology 195, 51-61, 1993
 A:Title: Sequence and genomic organization of Norwalk virus.
 A:Reference number: A37471; MUID:93303939; PMID:8391187
 A:Accession: B37471
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: genomic RNA
 A:Residues: 1-530 <JIA>
 A:Cross-references: GB:M87661; NID:gl061311; PIDN:AAB50466.1; PID:gl061313
 A>Note: sequence extracted from NCBI backbone (NCBIP:134157)
 C:Superfamily: human calicivirus capsid protein

Query Match 67.2%; Score 1946; DB 2; Length 530;
 Best Local Similarity 66.8%; Pred. No. 4.4e-137;
 Matches 365; Conservative 67; Mismatches 96; Indels 18; Gaps 7;

QY 1 MMWASKADPTNMDGTSGAGOLVPEANTAEPISEVPVAGATAAATAAGVNMIDPMMNMY 60
 DB 1 MMWASKDATSSVDGASGAGOLVPEVNASDPLANDPVPAGSATAVATAGVNPDPMLINN 60
 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGVGMKVKVLLAGNAFTA 120
 61 VQAPQGEFTISPNNTPGDVLFDLSLGLPHLNPFLSHLAQMYNGVGMKVKVLLAGNAFTA 120
 121 GKIIISICPPGPAQAQNTISQAATMFPHVIADRVLEPIEVPLEDVNVLPHND-NAPTM 179
 121 GKIIISICPPGFGSHNLIAQAATLFPHVIADRVLEPIEVPLEDVNVLPHNDNRNQQT 180
 180 RLVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFVLPVNPVEQTKPFSVNP 239
 181 RLVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFVLPVNPVEQTKPFSVNP 238
 240 NTLNSRVPSLIKMMVSRDHQMVQFQNGRVTLTGQLOGTTPTSASOLCKIRGSVFHAN 299
 239 SLSNSRAPLPISSIGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGT---S 294
 300 GGNGVNLTELDGSPYHAFESAPITGFDLGCEDHMEASPTQNTGDVTKQINVKQESA 359
 295 NGTVNLTELDGTPFHPPEGAPITGFDLGCEDHMEASPTQNTGDVTKQINVKQESA 351
 360 FAPHLGTTIADGLSDSVSNVNMIAKLGWSPVSDGHRGDVDPWVPIRYGSLTTEAAQLAP 419
 352 FVPHLGSIQANGIG----SGNVVGLVSWISPPSPSGSQVDLWKIPNYGSSITEATHLAP 407
 420 PIYPPGGEAIVFMSDFPIAHGTNGLSVPCITIQEFVTHFVNEQAPTRGEAALLHYLDP 479
 408 SVYPPGGEALVYFVSAPP---GPGAYNLPCLLPQBYIYTHFVSEQAPTWGDAALLHY 464
 480 DTHRLNLFGEFKLYPEGMTCVPNSSGTGPTLPINGVFFVSVWSRFPYQKPVGTA-G-C 538
 465 DTGRNLGEFKAYPDGFLICVFNAGSSGQQLPVGFFVSVWSRFPYQKPVGTAASSARG 524

QY 539 RLGIRR 544
 DB 525 RLGLRR 530

RESULT 3
 S60616
 capsid protein - human calicivirus (strain Melksham)
 C:Species: human calicivirus
 A:Variety: strain Melksham
 C:Date: 23-May-1997 #sequence_revision 23-May-1997 #text_change 28-Jul-2000
 C:Accession: S60616
 R:Green, S.M.; Lambdon, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
 Virus Res. 37, 271-283, 1995
 A:Title: Capsid diversity in small round-structured viruses: molecular characterization
 A:Reference number: S60615; MUID:96136658; PMID:8533462
 A:Accession: S60616
 A:Molecule type: genomic RNA
 A:Residues: 1-542 <GRE>
 A:Cross-references: EMBL:X81879; NID:g976077; PIDN:CAA57462.1; PID:g976079
 A:Experimental source: strain Melksham
 A>Note: it is uncertain whether Met-1 or Met-3 is the initiator
 C:Superfamily: human calicivirus capsid protein
 C:Keywords: capsid protein; coat protein

Query Match 42.2%; Score 1223; DB 2; Length 542;
 Best Local Similarity 47.2%; Pred. No. 3.2e-83;
 Matches 264; Conservative 73; Mismatches 188; Indels 34; Gaps 13;

QY 1 MMWASKADPTNMDGTSGAGOLVPEANTAEPISEVPVAGATAAATAAGVNMIDPMMNMY 60
 DB 1 MKWASNDAAAPSTDGAAG---LVPESNN-EVMALEPVAGAAALAAVPTGTQNIIDPWIRANF 56
 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGVGMKVKVLLAGNAFTA 120
 57 VQAPNCEFTVSPRNAPGCVLLNLELGPENLPYLALHARMYNGVAGMEVQVMLAGNAFTA 116
 121 GKIIISICPPGPAQAQNTISQAATMFPHVIADRVLEPIEVPLEDVNVLPH-NNDNAPTM 179
 117 GKLVFAAVPFPFVFNLSPOKITMFPHVIIDRVLEPIEVPLEDVNLSFFHYNQKDDPKM 176
 180 RLVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFVLPVNPVEQTKPFSVNP 239
 177 RIVAMLYTPLRNSG--SGDDVFTVSCRVLTRSPDFDFTVLPVPTVESKTKPFTLPILT 234
 240 NTLNSRVPSLIKMMVSRDHQMVQFQNGRVTLTGQLOGTTPTSASOLCKIRGSV---F 296
 235 GELNSRFPVPIDQMYTSPNEVISVQCONGRCTLDGELQGTTLQOVSGICAPKGEVTAHL 294
 297 HANGG-NGYNLTDLGSPYHAFES-PAPIGFPD-----LGECDHMEASPTQNTG 346
 295 HDNHLNNVTITNLNGSFFDPSEDIAPLGVDPDFQGRVFGVISQORDKQNAAGHSEPANRG 354
 347 -DVIKQINVKQESAFAPHLGTTIADGLSDSVSNVNMIAKLGWSPVSDGHRGDVDPWVPI 405
 355 HDVAVPTVTAQ---YTPKLGQIQITGWTQDLDLTQVNPVKF---TPVGLNTEHFNQWVP 408
 406 RYVSTLTTEAAQLAPPIYPPGGEAIVFMSDFPIAHGTNGLSVPCITIQEFVTHFVNEQA 465
 409 RYAGALNLTNLAPSVAPVFPGERLLFRSHLPKGGYGNPAIDCLLPQEWVQHYQEAA 468
 466 PTRYGAALLHYLDPDTHRLNLFGEFKLYPEGMTCVPNSSGTGPTLPINGVFFVSVWSR 525
 469 PSKSEVALVRYINPTDGRALFEAKLHRAGFTVTSNTS--APVVVPANGYFRFDSWVNQF 526
 526 YOLKPVGTAGPACRLGIRR 544
 527 YSLAPWGTGN-----GRRR 540

RESULT 4
 S40111
 capsid protein - human calicivirus (strain Bristol isolate B493)

C;Species: human calicivirus
A;Variety: strain Bristol isolate B493
C;Date: 25-Dec-1994 #sequence_revision 27-Feb-1997 #text_change 28-Jul-2000
C;Accession: S40111
R;Green, S.M.; Dingle, K.E.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
submitted to the EMBL Data Library, December 1993
A;Description: Human enteric caliciviridae: a new prevalent SRSV group defined by RNA-de
A;Reference number: S40111
A;Accession: S40111
A;Molecule type: genomic RNA
A;Residues: 1-539 <GRE>
A;Cross-references: EMBL:X76716; NID:9436410; PIDN:CAA54134.1; PID:9436411
A;Experimental source: human enteric calicivirus strain Bristol isolate B493
C;Superfamily: human calicivirus capsid protein
C;Keywords: capsid protein; coat protein

Query Match 39.8%; Score 1151.5; DB 2; Length 539;
Best Local Similarity 44.9%; Pred. No. 6.8e-78;
Matches 254; Conservative 78; Mismatches 185; Indels 49; Gaps 16;

Qy 1 MMASKADPTNMDGTSAGQLVPEANTAEPTISMEFPVAGAAATAATAGQNNMIDPWIMNYY 60
Db 1 MKMASNDANPS-DGS--AANLVPEVNN-EVMALEPVVGAATAAPVAGQNVDPWIRNPF 56

Qy 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNFPLSHLAQMYNGWGNMKVLLAGNAFTA 120
Db 57 VQAPQGEFTVSPRNAPGEILMSAPLGDPLNPLSHLSRMVNGYAGGFEVQVILAGNAFTA 116

Qy 121 GKIIISCIPCPGAQNTSIAQATPEPHVIAADVRLIEVLEPVEDVRLVLEH-NDNAPT 179
Db 117 GKVIAAAPPNPFTEGUSPSQVTPPHVIVDVRQLEPVLIPDPVNNFYHYNQANDSTL 176

Qy 180 RLVCWLYTPLRASGSSGSDTDFVIAGRVLTCPSPDFSLFLVPPNVEOKTKPFSPVNLPL 239
Db 177 KLIAMLYTPLRA--NNAGDDVFTVSCRVLTPSPDFDFILVPTVPSRTKPFVPLTV 234

Qy 240 NTLNSRVPSLISKMSVRDRHQVQFQNGRVLTDLQGLTTPTSASOLCKIRGSVFHAN 299
Db 235 EEMNSRPPIPLEKLYTGPSSAFVQVQNGRCTTGDVLLGTQLSANNICNFRGDVTHIA 294

Qy 300 GGNGY--NLTELDGSPYHAFES-PAPIGFPLD-----GECWHMEASPTTQPT 345
Db 295 GSHDYTNLASQNNNSYDPTBEIPAPLGTDPFVGKIQGLLTQTTRADGSTRAHKAT-VST 353

Qy 346 GDVIKQINVKQESAFAPHLGTIQ--ADGLSDVSVNTNMIAKLGWVSPVSDG--HRGDVDP 401
Db 354 GSV-----HFTPKLGSVQVTTDNDNFQAGQN--TKFTPGVVIQDGDHQNRPQQ 401

Qy 402 WVIPRYGSTLTEAQAQLAPPIYPPOGFEAIVFMSDFPIAHGTNGLSVCTTIPQBFVTHFV 461
Db 402 WLLPNYSGRTHNVHLAPAVAPTPEGEQLLFFRSTMPGCSGYNNMLDCLLPQEWVLHFY 461

Qy 462 NEQAPTRGEAALLHYLDPTHNLNLFKLYPEGFWTCVPNSSGTCGPQ--TLPINGVFVFV 519
Db 462 QEAAPAQSDVALLRFVNDPGRVLFCKLHKSGVITV---AHTGPVLDVLPPNGYPRFD 517

Qy 520 SWSRFYLKPVGTGAPACRIGRS 545
Db 518 SWNQFYTLAPMGNG-----TGRRA 538

RESULT 5
RRWRH
genome polyprotein - rabbit hemorrhagic disease virus
N;Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C;Species: rabbit hemorrhagic disease virus
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 29-May-1998
C;Accession: A41039
R;Meyers, G.; Wirblich, C.; Thiel, H.J.
Virology 184, 664-676, 1991
A;Title: Rabbit hemorrhagic disease virus--molecular cloning and nucleotide sequencing c
A;Reference number: A41039; MUID:91361557; PMID:1840711
A;Accession: A41039

A;Molecule type: genomic RNA
A;Residues: 1-2344 <MEY>
A;Cross-references: GB:M67473
C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C;Keywords: nucleotidyltransferase

Query Match 10.9%; Score 316; DB 1; Length 2344;
Best Local Similarity 28.3%; Pred. No. 9.6e-15;
Matches 132; Conservative 60; Mismatches 172; Indels 102; Gaps 21;

Qy 12 MDGTSAGQLVPEANTAEPTIS-----MEPVAGAATAA-----ATAG----- 47
Db 1766 MEGKARAAPGEEAGTATTASVPOTTTDGDMPGVVATTSVITAENSASIAATAGIGGPPQ 1825

Qy 48 QVNNMIDPWIMNYYVQAPQGEFTISPNNTPGDILFDLQGLPHLNFPLSHLAQMYNGWGNM 107
Db 1826 QVDOQETWRTNFYY---NDVFTMSVADAPGSILYTVQHSPPNNPFTAVLSQMYAGWAGGM 1882

Qy 108 KKVLLAGNAFTAGKIIISCIPCPGAQNTSIA---QATMPPHVIAADVRLIEVPLED 164
Db 1883 QFREIVAGSGVFGGLVRAVIPPQ-----IEIGFGLVRQFPFHVIDARSLEPVTITMPD 1937

Qy 165 VRNVLEH-NDNAPTMRMLVCMLYTPL--RASGSSGSDTDFVIAGRVLTCPSPDFSLFLV 221
Db 1938 LRPNMYHTPDGGLVPLTVLSVYNNLINPFGSIS-----AIQVTVETRPSEDFEFVIR 1992

Qy 222 PPNVEOKTKPFSPVNLPLNTLSNRVPSLISKMSVRDRHQVQFQ-----NGRVT 272
Db 1993 APS--SKTVDSISAPGLTT-----PVLGTGVDNRWNGQIVGLQVPVGFGSTCNRHN 2044

Qy 273 LDGLOQTTPTSASOLCKIRGSVFHANGGYNUTEL-----DGSFYHAFESP-----APIG 324
Db 2045 LNSGTGWSPPRFDDIDHRRGSASY-SGSNATNVLQFWYANAGS---AIDNPISQVAPDG 2100

Qy 325 FPDLEGCDWMEASPTTQNTGDVIKQINVKQESAFAPHLGTIQADGLSDVSVNTNMIAK 384
Db 2101 FPDMSFVFPNGPGLPAAGWGVFGGAIWNSN-----SGAPNVTTVQA-----YE 2142

Qy 385 LGWVSPVSDGHRGDVDPWVPIPRYGSTLTEAAQ-LAPPIYPGFGSEA 429
Db 2143 LGPAT---GAPGNLQP-----TTNTSQAQTVAKSIYAVVTGTA 2177

RESULT 6
A53982
capsid protein - European brown hare syndrome virus
C;Species: European brown hare syndrome virus
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 28-Jul-2000
C;Accession: A53982
R;Wirblich, C.; Meyers, G.; Ohlinger, V.F.; Capucci, L.; Eskens, U.; Thiel, H.
J. Virol. 68, 5164-5173, 1994
A;Title: European brown hare syndrome virus: relationship to rabbit hemorrhagic disease
A;Reference number: A53982; MUID:94309183; PMID:7518531
A;Accession: A53982
A;Status: Preliminary
A;Molecule type: mRNA
A;Residues: 1-576 <WIR>
A;Cross-references: GB:U09199
C;Superfamily: human calicivirus capsid protein

Query Match 10.9%; Score 315; DB 2; Length 576;
Best Local Similarity 26.2%; Pred. No. 1.5e-15;
Matches 126; Conservative 69; Mismatches 200; Indels 86; Gaps 20;

Qy 7 DAP-----TNMDGTSAGQLVPEANTAEPTISMEPVAGAAATAATAGQNNMIDP---WIMN 58
Db 8 DAPGTATTASVFGTGTDDGMDPGVVAVSDVTVADNVAAASVATAGTGGPPQASPOESWRVN 67

Qy 59 NYVQAPQGEFTISPNNTPGDILFDLQGLPHLNFPLSHLAQMYNGWGNMKVLLAGNAF 118
Db 68 FFY---NDVFTMSVTDAPGSLTYSVQHSPPNNPFTQVLSQMYAGWAGMQPFRFVAGSGI 124

Qy 119 TAGKIIISCIPPGFAAQ-NISIAQATMPPHVIAADVRLIEVPLEDVRNVLFH-NDNNA 176

Db 125 FGRLVCAIIIPGGIQGLEVRQ---PPhVIDARSLPEVTIIMPDLRPEMHTGDPG 181
Qy 177 PTMLRLVCMlyTFLRASGSSGTDPF-----VIAGRLVTCSPDSFSLFLVPPNVEQTKP 231
Db 182 LVPTLVSVYNNL-----INPFGTTSIAQVTVETRPSEDFEVLIRAPS---SKTVD 231
Qy 232 FSVNPLPLNTLSNRVPSLIKSMVMVRDGHQMVQFQ-----NGRVTLDGLOQTTP 282
Db 232 SVNFSWLLTT-----PVLTCAGSDNRWGAIPVGLQPVPGGFSNRRHNMNGSTYCWSS 285
Qy 283 TSASOLCKIRGSVFHANGGYNLTDLGSPYHAFESP-----APIGRDLGECDDHWEAS 338
Db 286 PRFDDIDHPSGNVSYPTGSATNTIETWYANAGTATTNPISNAPDGPDMGAIFP-----S 341
Qy 339 PTTQNTGDIKQINVKQESAPAPHLGTIQADGLSDSVNTMIKLGWSPVSDGHRGD 398
Db 342 GTT-IPTGAWVGFGQWNASNGTPYGVTVQA-----YELGF-----ANGAPSS 383
Qy 399 VDPWVIPRYGSTLTEAAQL-APPYIPPGFGE-----AIVFMSDFPIAHGTNGLSVPCITP 453
Db 384 IRP-----VTTTGAQLVAKSIYGAIVAQNSAGIIFLSKGMVS--TPGVAATTVP 434
Qy 454 Q 454
Db 435 Q 435
RESULT 7
genome polyprotein - rabbit hemorrhagic disease virus (isolate B889)
C:Species: rabbit hemorrhagic disease virus
A:Variety: isolate B89
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 18-Jun-1999
C:Accession: S55399
R:Rossi, C.
submitted to the EMBL Data Library, May 1995
A:Reference number: S55399
A:Accession: S55399
A:Molecule type: genomic RNA
A:Residues: 1-2344 <ROS>
A:Cross-references: EMBL:X87607; NID:g854640; PIDN:CAA60910.1; PID:g854641
A:Experimental source: isolate B89
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: polyprotein
Query Match 10.8%; Score 311.5; DB 2; Length 2344;
Best Local Similarity 27.4%; Pred. No. 2.1e-14;
Matches 127; Conservative 60; Mismatches 172; Indels 105; Gaps 20;
Qy 17 GAGQLVPEANTAEPISEPVAGAAAT-----AAATAG-----QV 49
Db 1768 GKARTAPOGEAAGTATTASVPGTTDGLDPGVATTSSVTAENSASATAGIGGPPQV 1827
Qy 50 NMIDPIMNNVQAPQGEFTISPNNTPGDILFDLQGLPHNLPFLSLHAQMYNGVMKV 109
Db 1828 DQQTWTNIFY---NDVFTWSVADAPGSLITYVQHSPPNNPFTAVLSQMYAGWAGMOP 1884
Qy 110 KVLLAGNAFTAGKIIISCIPGFAAQNISTA---QATMPFHVADRVRLPIEVPLEYR 166
Db 1885 RFIVAGSGVFGRLVAIVIPG-----IEIGPGLVQRQFPHVVIDAKSLPEVTTITPDLR 1939
Qy 167 NVLFH-NNDNAPTMRVLCMLYTPL--RASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPP 223
Db 1940 PNMVHTPGDGLVPTLVLSVYNNLINPFGSGTS-----AQVTVETRPSEDFEFVMIRAP 1994
Qy 224 NVEQTKRPFSPVNPFLNTLSNRVPSLIKSMVMVRDGHQMVQFQ-----NGRVTLD 274
Db 1995 S-SKTVDSISFAGLLTT-----PVLTVGVNDNRWNGQIVGLQFVPGGFSSTCRHWNLN 2046
Qy 275 QLOQTTPTTSALCKIRGVSFHANGGYNLTTEL---DGSPVHAFESP-----APIGFP 326
Db 2047 GSTYGMSSRPFADIDHRRGSASYP-GSNATNVQLQFWYANAGS---AVDNPISQVAPDGP 2102

Qy 327 DLGECDDHWEASPTTQENTGDVIRKQINVKQESAPAPHLGTIQADGLSDSVNTMIKLG 386
Db 2103 DMSFVFPNGPGIIPAGWVGFGAIWNSN-----SGAPNVTTVQA-----YELG 2144
Qy 387 WVSVPVSDGHRGDVDPWVIPRYGSTLTEAAQ-LAPPIYPPGFGEA 429
Db 2145 FAT---GAPGNLQP-----TTNTSGAQTVAKSIYAVVTGTA 2177
RESULT 8
S64740
genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)
N:Contains: VP60 protein
C:Species: rabbit hemorrhagic disease virus
A:Variety: isolate AST/89
C:Date: 12-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Jun-2000
C:Accession: S64740; S49018; S49018; S49018
R:Casais, R.; Martin-Alonso, J.; Boga, J.; Parra, F.
submitted to the EMBL Data Library, May 1995
A:Description: Genomic organization of rabbit hemorrhagic disease virus determined by di
A:Reference number: S64740
A:Accession: S64740
A:Molecule type: genomic RNA
A:Residues: 1-2344 <CAS>
A:Cross-references: EMBL:Z49271; NID:g1182032; PIDN:CAA89265.1; PID:g1150552
A:Experimental source: isolate AST/89
R:Boga, J.; Casais, R.; Marin, M.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra, F.
submitted to the EMBL Data Library, July 1993
A:Description: Molecular cloning, sequence and expression of the capsid protein gene fro
A:Reference number: S46944
A:Accession: S46944
A:Molecule type: genomic RNA
A:Residues: 1650-2344 <BOG>
A:Cross-references: EMBL:Z24757; NID:g515622; PIDN:CAA80881.1; PID:g515623
A:Experimental source: isolate AST/89
R:Parra, F.; Boga, J.A.; Marin, M.S.; Casais, R.
Virus Res. 27, 219-228, 1993
A:Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus suppo
A:Reference number: S49018; MUID:93255896; PMID:8488721
A:Accession: S49018
A:Molecule type: genomic RNA
A:Residues: 1650-1796 <PAR>
A:Cross-references: EMBL:Z24757
A:Experimental source: isolate AST/89
A:Accession: S65012
A:Molecule type: protein
A:Residues: 1767-1779; 1875-1877, 'X', 1879-1881; 1936-1938, 'X', 1940-1941 <PAW>
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: blocked amino end; polyprotein
Query Match 10.6%; Score 307.5; DB 2; Length 2344;
Best Local Similarity 26.3%; Pred. No. 4.1e-14;
Matches 149; Conservative 67; Mismatches 212; Indels 139; Gaps 25;
Qy 12 MDG---TSGAQLVPEANTAE---PISMEPVAGAAATAA-----ATAG----- 47
Db 1766 MEGKARTAPOGEAAGTATTASVPGTTDGLDPGVATTSSVTAENSASATAGIGGPPQ 1825
Qy 48 QVNMIDPIMNNVQAPQGEFTISPNNTPGDILFDLQGLPHNLPFLSLHAQMYNGVMKV 107
Db 1826 QVDDQETWRTNIFY---NDVFTWSVADAPGSLITYVQHSPPNNPFTAVLSQMYAGWAGM 1882
Qy 108 KVKVLLAGNAFTAGKIIISCIPGFAAQNISTA---QATMPFHVADRVRLPIEVPLEYR 164
Db 1883 QPRFIVAGVFGGRLVAIVIPG-----IEIGPGLVQRQFPHVVIDAKSLPEVTTITP 1937
Qy 165 VNRVLFH-NNDNAPTMRVLCMLYTPL--RASGSSSGTDPFVIAGRVLTCPSPDFSLFLV 221
Db 1938 LRPNMVHTPGDGLVPTLVLSVYNNLINPFGSGTS-----AQVTVETRPSEDFEFVMIR 1992
Qy 222 PPNVEQTKRPFSPVNPFLNTLSNRVPSLIKSMVMVRDGHQMVQFQ-----NGRVT 272

```
Db 1993 APS--SKTVDISAPAGLLTT-----PVLTVGNDNRWNGQIVGLQVPVPGGFCSTCRHWN 2044
Qy 273 LDGQLQGTTPTSASQCKIRGSVHFANGNGVNYTEL-----DGSPYHAFESP-----APIG 324
Db 2045 LNSTYGVSSPRFGDIGHRGSASYP-GNNATNVLQFWYANAGS---AIDNPISQVAPDG 2100
Qy 325 FPDLGECDDHMEASPTTQFNTGDIKQINVKQESAFAPHLGTIQADGLSDVSVNTNIAK 384
Db 2101 FPDMSFVDFNGFIPAGWVGFGAIWNSN-----SGAPNVTVQA-----YE 2142
Qy 385 LGWVSPVSDGHRGDVDPWVPIPRYGSTLTERAAQ-LAPPIYPGFGGAI-----VFFWSDPPI 439
Db 2143 LGPAT-----GAPGNLQ-----TTTSGQTWAKSIYAVVTGTAQNPAGLFVNASGVI 2191
Qy 440 AHGNTGLSVPCCTIQPFVTHFVNEQAPTRGEGALLHYLDPDTHRNGLGFKLYPGEFTCV 499
Db 2192 STPSANAITYPQPDRIVT-----TPGTAAAPVGNKTPIMFASVW 2232
Qy 500 PNSS-----GTGPQTLPI 512
Db 2233 RRTGDVNATAGSANGTQYGTGSQLPV 2259

RESULT 9
A:848562
coat protein - San Miguel sea lion virus (serotype 1)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: A48562
R:Neill, J.D.
Virus Res. 24, 211-222, 1992
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lions.
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Accession: A48562
A:Molecule type: genomic RNA
A:Residues: 1-702 <NEI>
A:Cross-references: GB:M87481; NID:G334882; PIDN:AAA16217.1; PID:G334884
A:Note: sequence extracted from NCBI backbone (NCBIN:113564, NCBIPI:113565)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.4%; Score 300.5; DB 1; Length 702;
Best Local Similarity 25.2%; Pred. No. 2.4e-14;
Matches 137; Conservative 66; Mismatches 209; Indels 131; Gaps 25;

Qy 13 DGTSGAGQLVPEANTAPISMEPVAGATAATAAGQVNMID-PM-INNNYVQAPQGBFTI 70
Db 154 DPGGADIVTBEQGTVVOQQVPQAQSALTTLTAAASTGKTVDCEWTFPSYHTA-----VNW 209
Qy 71 SPNNTPGDILFDLQGHNLNPLSHLAQMYNGVGNMKVLLAGNAFTAGKIIISCIPP 130
Db 210 STTEAQGKILFSRALSPELNPLRHISLSYSTWSGGIDVREFTVSGVFGKLAALIVPP 269
Qy 131 GFAAQNTISIAQTAFPHVIAVRLVLEIEVPLEDVRNVLPHNNDAFTMLRVCMLYTPL- 189
Db 270 GI--EPVESPTMLOYPHVLFPARQTEPVIFIPDIRKTLHSMDDTDTTLVIMVYNELI 327
Qy 190 -----RASGSSGTDPPVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVN--LPPLNT 241
Db 328 NPYEQSEPKSSCSIT-----VETRPSSDFTFSLKPPG--SLLKHGSIPLSPRNS 377
Qy 242 --LNSRVPSLIKSMVSRDHGMVQFQNGR-----VTLD----- 274
Db 378 RHMGNRNMWSTIDGFV-----QPRVFQSNRHFDFDSTTGWSTPYIPIEVLTKLDRG 432
Qy 275 GOLQGTTPTSASQL-----CKIRGSVPHANGNGVNYTEL-----DGSPYHAFESPA 321
Db 433 GQYFKVTDTEKSLVPGLPDGHGPDPTTIPMTASNGNVDYTVAEIRITNNGTHFGKFTMG 492
Qy 322 PI-----GFPDLGECDDHMEASPTTQFNTGDIKQINVKQESAFAP-----HLGTIOA 369
```

```
Db 493 NLTTKVGKSNLGET---QOTSRTLFAVG-----NYKOQNTINPTTHKITSNSLVVYDA 543
Qy 370 DGLSDSVSVNT-----NMIAGLQWV-----SPVSDGHRGDVDPWVPIPRYGSTLTERAAQLA--P 419
Db 544 NVVSAATAKTITTTWISTHSLGVLVDESVPV-----GSDSTKVVRIATLP 587
Qy 420 PIYPGPGGAIIVFPMSPFPIAHGNTGLSVPCCTIQPFVTHFVNEQAPTRGEGALLHY-LD 478
Db 588 EAFVNG-GNPFVFTNKIQIGH-----FDRAHTKCFNSQVLMTSOKLAENHYTLP 636
Qy 479 PDT 481
Db 637 PDS 639

RESULT 10
JQ2354
capsid protein - feline calicivirus (strain NADC)
C:Species: feline calicivirus
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: JQ2354
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.
J.Gen. Virol. 74, 2519-2524, 1993
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable regions.
A:Reference number: JQ2354; MUID:94065683; PMID:7504075
A:Accession: JQ2354
A:Molecule type: mRNA
A:Residues: 1-668 <SEA>
A:Cross-references: GB:I09718; NID:G305104; PIDN:AAA16485.1; PID:G305105
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein

Query Match 9.6%; Score 277.5; DB 2; Length 668;
Best Local Similarity 23.3%; Pred. No. 1.2e-12;
Matches 124; Conservative 75; Mismatches 178; Indels 155; Gaps 24;

Qy 34 EPVAAATAATAAGQVNMIDPMINNYVQAPQGBFTI-----SPNNTPGDILFDLQGLP 87
Db 145 EPSAQMSTAADMATGKGVDSW-----EAFPSFHTSVNMWSTSETQGIKLPKQSLGP 195
Qy 88 HLANPFLSHLAQMYNGVGNMKVLLAGNAFTAGKIIISCIPPGF-AAQNTISIAQTAFNP 146
Db 196 LLNPYLEHLKLYVAVSGSVSEVFSISGSGVFGKLAALVVPVGVDPVQSTMLQ---Y 252
Qy 147 HVIADVRLVLEPIEVLPLEDVRNVLPHNNDAFTMLRVCMLY-----TPLRASGSSGTDPPV 202
Db 253 HVLFDAQVDPVIFSIPLASTLYHLMPTDITSLVIMVYNDLINPYNANDSNSSGC---I 309
Qy 203 IAGRVLTCPSPDFSLFLVPP----- 223
Db 310 VT--VETKPGDPKPKFHLKPPGSMLTGHSVPSDLIPKSSSLWIGNRYWSDITDFVVRPFV 367
Qy 224 -----NVQKTKPFPVPLN-PLN-TLSNRVPSLIKSMVSRDHGMVQFQNGRVTLTD 274
Db 368 FOANRHFNFQETAGWSAPRFPITITISKSGKL--GIGVATDY-----IVP 414
Qy 275 GOLQGTTPTSASQLCKIRGSVPHANGNGVNYTELSDGSPYHAFE-SPAPIGPDLGECDD- 332
Db 415 GIPDGWPDITIAEDLTTPAGD-YAITSNGNDIT--TGSEYDSTVEIKNTNFRGMYICGS 471
Qy 333 ----WHMEASPTTQF-----NTGDVIKQINVKQESAFAPHLGTIQADGLSD 374
Db 472 LQRAWGDKKISNTAFITTAKEGNKIRPSNTIDMTK-IADVQDT---HVG-----EE 519
Qy 375 VSVNTNIAKLGWVSPVSDGHRGDVDP-----WVPIPRYGSTLTERAAQLAPPIYPGFGGAI 430
Db 520 VQTSDDALLALTYGIGEQAIGSDRDRVVRISVLPEVGAR-----GNHP 564
Qy 431 VFFWSDPPIAHGNTGLSVPCCTIQPFVTHFVNEQAPTRGEGALLHY-LDPDT 481
Db 565 IFYKNSIKUGVIRSIDV-----FNSQILHTRSQLSLNHYLLSPDS 605
```

VCWVFC coat protein - feline calicivirus (strain Japanese F4)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C:Accession: B40481
R:Tohya, Y.; Taniguchi, Y.; Takahashi, E.; Utagawa, E.; Takeda, N.; Miyamura, K.; Yamaza
Virology 183, 810-814, 1991
A:Title: Sequence analysis of the 3'-end of feline calicivirus genome.
A:Reference number: A40481; MUID:91306470; PMID:1853578
A:Accession: B40481
A:Molecule type: Genomic RNA
A:Residues: 1-668 <TOH>
A:Cross-references: GB:D90357; NID:g221264; PIDN:BAA14371.1; PID:g221266
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,301,304,399,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:177,301,304,399,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.5%; Score 274; DB 1; Length 668;
Best Local Similarity 31.0%; Pred. No. 2.1e-12;
Matches 74; Conservative 39; Mismatches 82; Indels 44; Gaps 9;

Qy 23 PEANTA-EPISMEPVAGATAAATAGQVNMIDPWIMNNVQAPQGEFTI-----SPNNT 75
Db 133 PEQGTAVGVIAEPPSAQMSAADMASGKSVDSW-----EAFFSHTSVNWSSTST 183
Qy 76 PGDILFDLQGLHNPFLSHLAQMYNGWGNMKVYLLAGNAFTAGKIISCIPPGF-AA 134
Db 184 QGKILFKQSLGPIILNPFYLEHSLKLYVWSSGSEVRSISGSGVFGKLAIVVPPGVDPV 243
Qy 135 QNTISIAQATMFPFHVADRVLEPIEVLNVLPHNDNAPTMELVCMY-----TPLR 190
Db 244 QSTSMQLQ---YPHVLFDAQVPEVIFCLPDLRSTLYHLMSTDTTSLVIMVNDLINPYA 300
Qy 191 ASSGSSGTDGFVIAAGRVLTCPSPDFSLFLVPPNVQKTKPFSPVPLNTLSNSRVPS 249
Db 301 NDSNSSGC-----IVT--VETKPGDPDFHLLKPPG-----SVLTGHSIPS 339

RESULT 13
VCWVFF coat protein - feline calicivirus (strain CFI/68 FIV)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Aug-1999
C:Accession: A40507; B40507; T09246
R:Neill, J.D.; Reardon, I.M.; Heinrikson, R.L.
J. Virol. 65, 5440-5447, 1991
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline caliciv
A:Reference number: A40507; MUID:91374597; PMID:1716692
A:Accession: A40507
A:Molecule type: Genomic RNA
A:Residues: 1-668 <NET>
A:Cross-references: GB:M32819; NID:g323874; PIDN:AAA42925.1; PID:g323875
A:Accession: B40507
A:Molecule type: protein
A:Residues: 373-379; 403-419; 481-489; 560-566 <NE2>
R:Neill, J.D.
submitted to the EMBL Data Library, April 1998
A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.
A:Reference number: Z1626
A:Accession: T09246
A>Status: preliminary, translated from GB/EMBL/DBJ
A:Molecule type: Genomic RNA
A:Residues: 1-668 <NE3>
A:Cross-references: EMBL:U13992; NID:g3056875; PIDN:AAC13993.1; PID:g537256
A:Experimental source: strain CFI/68 FIV
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.3%; Score 269; DB 1; Length 668;
Best Local Similarity 31.0%; Pred. No. 5e-12;

VCWVFF coat protein - feline calicivirus (strain F9)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C>Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 28-Jul-2000
C:Accession: B43382; C45538; FQ0407; S23702
R:Carter, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.; Turner, P.C.
Virology 190, 443-448, 1992
A:Title: The complete nucleotide sequence of a feline calicivirus.
A:Reference number: A43382; MUID:92410623; PMID:1529544
A:Accession: B43382
A:Molecule type: Genomic RNA
A:Residues: 1-671 <CAR1>
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879
R:Carter, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.
Arch. Virol. 122, 223-235, 1992
A:Title: Identification and sequence determination of the capsid protein gene of feline
A:Reference number: A45538; MUID:92117861; PMID:1731695
A:Accession: C45538
A:Molecule type: Genomic RNA; protein
A:Residues: 1-671 <CAR2>
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879
A:Experimental source: strain F9
A:Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)
R:Quiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.
J. Gen. Virol. 73, 2429-2433, 1992
A:Title: The cloning, sequencing and expression of a major antigenic region from the fel
A:Reference number: FQ0407; MUID:93019069; PMID:1402818
A:Accession: PQ0407
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 217-266, 'ST', 269-336, 'V', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>
A:Cross-references: PIDN:AB23553.1; PID:g257083
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.5%; Score 276; DB 1; Length 671;
Best Local Similarity 26.4%; Pred. No. 1.5e-12;
Matches 102; Conservative 54; Mismatches 128; Indels 102; Gaps 17;

Qy 13 DGTSGAGQLVPEANT-AEPISEMPVAGATAAATAGQVNMIDPWIMNNVQAPQGEFTI- 70
Db 127 DGSITA----PEQGTWGGVIAEPPSAQMSAADMATGKSVDSW-----EAFFSPH 173
Qy 71 -----SPNNTPGDILFDLQGLHNPFLSHLAQMYNGWGNMKVYLLAGNAFTAGKI 125
Db 174 TSNWSTSETQKILFKQSLGPIILNPFYLEHSLKLYVWSSGSEVRSISGSGVFGKLA 233
Qy 126 SCIPPGF-AAQNTISIAQATMFPFHVADRVLEPIEVLNVLPHNDNAPTMELVCM 184
Db 234 IVVPPGVDPVQSTSMQLQ---YPHVLFDAQVPEVIFCLPDLRSTLYHLMSTDTTSLVIM 290
Qy 185 LY-----TPLRASSGSSGTDGFVIAAGRVLTCPSPDFSLFLVPPNVQKTKPFSPVPLPLN 240
Db 291 VYNDLINPYANDANSSGC-----IVT--VETKPGDPDFHLLKPPG-----SMLTHG 335
Qy 241 TLSNSRVPSPLIKSMVMVRDHQVQFQNGRVTLTGDLQGLQGTTPTSASOLCKIRGSVFHANG 300
Db 336 SIPSDLIPKTSLSLIGNRYWSDITDF-----VIRPFVQANR 372
Qy 301 GNGYN-----LTELDCGSPHYA-----FESPA-PIGFPPDL---GEC--- 331
Db 373 HFDENQETAGASTFRFRPISVTITTEONGAKLIGVATDIIVPGIDPGWPDITTPGELIPA 432
Qy 332 -DWMH-----EASPTTQFNQGDVNIK 350
Db 433 GDYAITNGTGNDDITATGYDTADIIK 458

RESULT 12

Matches 74; Conservative 38; Mismatches 83; Indels 44; Gaps 9;

Qy 23 PEANT-AEPISEMPVAGATAAATAGQVNMIDPWIMNYYVQAPQGEFTI-----SPNNT 75
Db 133 PEOGTWVGVIAPNAQNAQSTADMATGKSDVSEW-----EAFPSFHTSVNMSTSET 183
Qy 76 PGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAPTAKKIIISCIPPGF-AA 134
Db 184 QGKILFKOSGLPLNPLYTHLAKLYVAMSGSDVRFSGVSGVFGKLAALVPPGIDPV 243
Qy 135 QNISTAQATMPHVIADVRVLEPIEVPLEDVRNVLPHNDNAPTMRVLCMLY----TPLR 190
Db 244 QSTWMLQ---YPHVLFDAQVPEVIFSPDURSTLYHLMSDITDTSVLYMINDLINPYA 300
Qy 191 ASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVQKTPKPSVPLNLTLSNSRVPS 249
Db 301 NDSNSSGC---IVT--VETKGPDPFKHLLKPPG-----SMLTHGSIIPS 339

RESULT 14
C48562
coat protein - San Miguel sea lion virus (serotype 4)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: C48562
R:Neill, J.D.
Virus Res. 24, 211-222, 1992
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lions.
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Accession: C48562
A:Molecule type: Genomic RNA
A:Residues: 1-703 <NEI>
A:Cross-references: GB:M37482; NID:g334886; PIDN:AAAL6220.1; PID:g334888
A:Note: sequence extracted from NCBI backbone (NCBIF:113567)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:89,208,329,463,482/Binding site: carbohydrate (Asn) #status predicted

Query Match 9.1%; Score 264; DB 1; Length 703;
Best Local Similarity 27.4%; Pred. No. 1.3e-11;
Matches 104; Conservative 53; Mismatches 181; Indels 42; Gaps 17;

Qy 13 DGTSGAGOLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNYYVQAPQGEFTISP 72
Db 154 DPGSAEIVTEOGTVVQQQAPAPTALATATATGKSVQEWTFP--SYHTSINWST 211
Qy 73 NNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAPTAKKIIISCIPPGF 132
Db 212 VESQKILYSQALNPSINPYLDHIAKLYSTWSGGIDVRFTVSGVFGKLAALLVPPGV 271
Qy 133 -AAQNISTAQATMPHVIADVRVLEPIEVPLEDVRNVLPHNDNAPTMRVLCMLYTP 191
Db 272 EPIESVSMLO---YPHVLFDAQVPEVIFTPIDIRKTLFHSMDETDTTKLVINPY----E 324
Qy 192 SGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVQKTPKPSVPLNLTLSNSRV 247
Db 325 NGVENKTTCSI---VETRPADFTFALLKPPG--SLIKGSIPLDIPRNSAHWMGNRW 379
Qy 248 PSLIKSMVMVRDHQGMVQFQNGR-VTLDGQLQG-TTPTSASQLCKIRGVSFHANGGNG-- 303
Db 380 WSTISGFSV-----QPRVFQSNRHFDFDSTTTGWSGTTPYVVFIEIKQGV----GSNNKW 430
Qy 304 YNLTELDGSPHAFESAPIGFPLGECDWMEASPTT-QNTGVDVK--QINVKQESAF 360
Db 431 FHVIDTD-----KALVPGIPDGWPTTIPD-----ETKATNGNFSYGESVRAGSTTIKPNENS 483
Qy 361 APHLGTIQADGLSDVSVNNTN 380
Db 484 THFKGTTCGLTSTVEIPEN 503

RESULT 15
JQ2356
capsid protein - feline calicivirus (strain KCD)
C:Species: feline calicivirus
C>Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: JQ2356
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable
A:Reference number: JQ2354; MUID:94065683; PMID:7504075
A:Accession: JQ2356
A:Molecule type: mRNA
A:Residues: 1-668 <SEA>
A:Cross-references: GB:L09719; NID:g305107; PIDN:AAAL6487.1; PID:g305108
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein

Query Match 9.0%; Score 261.5; DB 2; Length 668;
Best Local Similarity 31.1%; Pred. No. 1.8e-11;
Matches 76; Conservative 37; Mismatches 86; Indels 45; Gaps 10;

Qy 19 GOLV-PEANT-AEPISEMPVAGATAAATAGQVNMIDPWIMNYYVQAPQGEFTI----- 70
Db 128 GSIVTPEOGTWVGVIAPNAQNAQSTADMATGKSDVSEW-----EAFPSFHTSVNM 178
Qy 71 SPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAPTAKKIIISCIPP 130
Db 179 STSETQKILFKOSGLPLNPLYTHLAKLYVAMSGSIEVRFSISGSGVFGKLAALVVP 238
Qy 131 GF-AAQNISTAQATMPHVIADVRVLEPIEVPLEDVRNVLPHNDNAPTMRVLCMLY--- 186
Db 239 GVDPVQSTSMLO---YPHVLFDAQVPEVIFSPIDLRSTLYHLMSDITDTSVLMVAYNDL 295
Qy 187 -TPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVQKTPKPSVPLNLTLSNS 245
Db 296 INPVANDSNSSGC---IVT--VETKGSDFRHLKPPG-----SVLTHG 335
Qy 246 RVPS 249
Db 336 SVPS 339

Search completed: June 1, 2004, 13:55:17
Job time : 12.5159 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:34:56 ; Search time 6.58389 Seconds
(without alignments)
4310.252 Million cell updates/sec

Title: US-09-926-799-1

Perfect score: 2896

Sequence: 1 MMASKDAPTNDGTSGAGQ.....YQLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1974.5	68.2	546	1 COAT_SOUV3	Q04542 southampton
2	1148.5	39.7	539	1 COAT_LORDV	P34635 lordsdale v
3	316	10.9	2344	1 POLN_RHDV	P27410 rabbit hemo
4	300.5	10.4	702	1 COAT_SMSV1	P36284 san miguel
5	289	10.0	2208	1 POLN_MANCV	Q69014 manchester
6	276	9.5	671	1 COAT_FCVP9	P27405 feline cali
7	274	9.5	668	1 COAT_FCVP4	P27405 feline cali
8	269	9.3	668	1 COAT_FCVC6	P27404 feline cali
9	264	9.1	703	1 COAT_SMSV4	P36285 san miguel
10	146.5	5.1	2205	1 POLG_POL2W	P23069 p genome po
11	145.5	5.0	2207	1 POLG_POL2L	P06210 p genome po
12	145	5.0	2206	1 POLG_POL3L	P03302 poliovirus
13	143.5	5.0	2194	1 POLG_HE701	P25537 h genome po
14	139.5	4.8	2206	1 POLG_POL32	P06209 poliovirus
15	137.5	4.7	2209	1 POLG_POL1S	P03301 p genome po
16	136.5	4.7	2208	1 POLH_POL1M	P03300 p genome po
17	135.5	4.7	2206	1 POLG_POL1M	P03299 p genome po
18	133.5	4.6	2179	1 POLG_HRV14	P03303 h genome po
19	124.5	4.3	855	1 POLG_HRV3	Q82081 human rhino
20	124	4.3	2214	1 POLG_CXA24	P36290 c genome po
21	123.5	4.3	2175	1 POLG_BOVEV	P12915 b genome po
22	119.5	4.1	2196	1 POLG_EC05N	Q9v1j1 e genome po
23	118.5	4.1	2194	1 POLG_EC30B	Q9v1j1 e genome po
24	117.5	4.1	2185	1 POLG_CXB5P	Q03053 c genome po
25	117	4.0	788	1 BCSB_XANAC	P58933 xanthomonas
26	117	4.0	2183	1 POLG_CXB4E	Q86887 c genome po
27	117	4.0	2183	1 POLG_CXB4J	P08292 c genome po
28	116.5	4.0	2206	1 POLG_CXA21	P22055 c genome po
29	115	4.0	2185	1 POLG_SVDVH	P16604 s genome po
30	115	4.0	2185	1 POLG_SVDVU	P13900 s genome po
31	113.5	3.9	2164	1 POLG_HRV89	P07210 h genome po
32	112.5	3.9	2184	1 POLG_EC01F	O91734 e genome po
33	112	3.9	2195	1 POLG_EC11G	P29813 e genome po

RESULT 1

COAT_SOUV3 ID COAT_SOUV3 STANDARD; PRT; 546 AA.

AC Q04542; 01-OCT-1994 (Rel. 30, Created)

DT 01-OCT-1994 (Rel. 30, Last sequence update)

DT 01-OCT-1994 (Rel. 30, Last annotation update)

DE Coat protein (Capsid protein).

OS Southampton virus (serotype 3).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC ~~Nonoviridae~~

OK NCBI_TaxID=37129;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93142023; PubMed=8380940;

RA Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;

RT "Sequence and genome organization of a human small round-structured

RT (Norwalk-like) virus."

RL ~~Sequence=2597516-539(49993)~~

CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).

CC -----

DR EMBL; L07418; AAA92984.1; --

DR PIR; B37491; B37491.

DR InterPro; IPR004005; Calici_coat.

DR InterPro; IPR008975; Viral_Cap_coat.

DR Pfam; PF00915; Calici_coat; 1.

KW Coat protein; Glycoprotein.

FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 546 AA; 58774 MW; C82B2A85AD4B05EA CRC64;

Query Match 68.2%; Score 1974.5; DB 1; Length 546;

Best Local Similarity 66.6%; Pred. No. 2e-141;

Matches 367; Conservative 75; Mismatches 96; Indels 13; Gaps 7;

QY 1 MMASKDAPTNDGTSGAGQLVPEANTAEPTSMSPVAGAAATAAGVNMIDPVMNNY 60

DB 1 MMASKDAPQSGASGAGQLVPEVNTADPLPMPFVAGPTTAVATAGVNMIDPVMNNF 60

QY 61 VQAPQGEFTTSPNTPGDILFDLQGLPHLPPLSHLAQMYNGWGNMKVYLLAGNAFTA 120

DB 61 VQSPQGEFTTSPNTPGDILFDLQGLPHLPPLSHLSQMYNGWGNMKVRVILLAGNAPSA 120

QY 121 GKIIISICPPGFAQNISIAQATMPHVIADVRVLEIEVPLEDVRNVLFHNNNDNAPTMR 180

DB 121 GKIIIVCCVPPGFTSSSITIAQATLPHVIADVRTLEIEMPLEDVRNVLYHTNDNQPTMR 180

QY 121 GKIIIVCCVPPGFTSSSITIAQATLPHVIADVRTLEIEMPLEDVRNVLYHTNDNQPTMR 180

DB 121 GKIIIVCCVPPGFTSSSITIAQATLPHVIADVRTLEIEMPLEDVRNVLYHTNDNQPTMR 180

```

RESULT 2
COAT_LORDV          COAT_LORDV          STANDARD;          PRT;          539 AA.
AC PS4635;
01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
Coat protein (Capsid protein).
Lordsdale virus (Human enteric calicivirus).
Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
Nobovirus.
NCBI_TaxID=82658;
[1]
SEQUENCE FROM N.A.
MEDLINE=96005060; PubMed=7561776;
Dingle K.E., Lambden P.R., Caul E.O., Clarke I.N.;
"Human enteric Caliciviridae: the complete genome sequence and
expression of virus-like particles from a genetic group II small
round structured virus";
J. Gen. Virol. 76:2349-2355 (1995).
RL
-!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
-----
This SWISS-PROT entry is copyright. it is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement)
or send an email to license@isb-sib.ch).
-----
EMBL; X86557; CAA60255.1; -
InterPro; IPR004005; Calici_coat.
InterPro; IPR008975; Viral_cap_coat.
Pfam; PF00915; Calici_coat; 1.
Coat protein; Glycoprotein.
CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 531 531 N-LINKED (GLCNAC. .) (POTENTIAL).
FT FT SEQUENCE 539 AA; 58775 MW; 5E5C63E7F2C5FD21 CRC64;
Query Match 39.7%; Score 1148.5; DB 1; Length 539;
Near Local Similarity 44.9%; Pred. No. 4.1e-79;

```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M67473; AAA47285.1; -;
 PIR; A41039; RWRWRH.
 PDB; 1KHV; 16-JAN-02.
 PDB; 1KHW; 16-JAN-02.
 MEROPS; C24.001; -;
 InterPro; IPR003593; AAA ATPase.
 InterPro; IPR004005; Calici coat.
 InterPro; IPR004004; Calici_pol_hel.
 InterPro; IPR000317; Peptidase C24.
 InterPro; IPR006005; RNA helicase.
 InterPro; IPR007095; RNA pol DS PS.
 InterPro; IPR001205; RNA pol 3D.
 InterPro; IPR007094; RNA pol_Psvir.
 InterPro; IPR008975; Viral_cap_coat.
 Pfam; PF00915; Calici_coat; 1.
 Pfam; PF03510; Peptidase C24; 1.
 Pfam; PF00680; RNA dep RNA_pol; 1.
 Pfam; PF00910; RNA_helicase; 1.
 PRINTS; PR00916; 2CENDOPTASE.
 PRINTS; PR00918; CALICIVIRUSNS.
 SMART; SM00382; AAA; 1.
 PolyProtein; Transferase; RNA-directed RNA polymerase; Hydrolase;
 Thiol protease; Helicase; ATP-binding; Coat protein; 3D-structure.
 CHAIN 7508 7707
 CHAIN 1109 1251
 CHAIN ? 2344
 NP_BIND 522 529
 ACT_SITE 1212 1212
 ACT_SITE 1227 1227
 SEQUENCE 2344 AA; 257066 MW; 1454C248F81E9212 CRC64;

Query Match 10.9%; Score 316; DB 1; Length 2344;
 Best Local Similarity 28.3%; Pred. No. 2e-15;
 Matches 132; Conservative 60; Mismatches 172; Indels 102; Gaps 21;
 Qy 12 MDGTSAGQLVPEANTAPIS-----MEPVAGATAA-----ATAG----- 47
 Db 1766 MEGKARAAPGAEAGTATTASVPGTTTDMGDPGVVATTSVITAENSSASITAGIGPPQ 1825
 Qy 48 QVNMIDPMIMNYVQAPQGETISNNTPGDLPLGLPHNPLSLAQMNGWGNM 107
 Db 1826 QVDOQETWRTHFY---NDVFTWSVADAPGSLITYQHSQNNPPTAVLSQMYAGWAGM 1882
 Qy 108 KVKVLLAGNAFTAGKIISCIPPGFAAQNI SIA---QATMFPFHVIADVRLVLEPIEVPLED 164
 Db 1883 QFRFVAGSGVFGGLVRAVIPPQ-----IEIGPGLVVRQFPFHVIDARSLPVTITMPD 1937
 Qy 165 VRNVLFH--NNNAPTRMLVCMLYTPL--RAGSSSGTDPFVIAGRVLTCPSPDFSLV 221
 Db 1938 LRPNMYHTGDPGLVTLVLSVYNNLINPFGGSTS-----AIQVTVTRPSEDFEVMIR 1992
 Qy 222 PPNVEQTKPFSVNLPLNTLSNRVPSLISKSMVSRDHGMVQVQF-----NGRVT 272
 Db 1993 APS---SKTVDSISPAGLLT-----PVLTVGNDNRNGQVLGLQVPVGFSTCNRHN 2044
 Qy 273 LDGLOQTTPTSASOLCKIRGSVFHANGGYNLTTEL-----DGSPYHAFESP-----APIG 324
 Db 2045 LNSGTYGWSSPRFGDIDHRRGSASY--SGSNATNVLFQWYANAGS---AIDNPIQVAPDG 2100
 Qy 325 PFDLGECMDHMEASPTTQNTGDIKQINVKQESAFAPHLGTIQADGLSDSVNNTMIATK 384
 Db 2101 PFDMSFVFNPGIPGIPAAQWVGFGALWNSN-----SGAPNVTTVQA-----YE 2142
 Qy 385 LGWVSPVSDGHRGDVDPWVI PRYGSTLTAAQ--LAPPIYPPFGFEA 429
 Db 2143 LGPAT---GAPGNLQ-----TTNTSGATVAKSIYAVVTGTA 2177

RESULT 4
 COAT_SMSV1 STANDARD; PRT; 702 AA.
 ID COAT_SMSV1 STANDARD; PRT; 702 AA.
 AC P36284;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Coat protein (Capsid protein).
 OS San Miguel sea lion virus (serotype 1) (SMSV 1).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus
 OX NCBI_TaxID=36406;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92410750; PubMed=1529644;
 RA Neill J.D.;
 RT "Nucleotide sequence of the capsid protein gene of two serotypes of
 San Miguel sea lion virus: identification of conserved and non-
 conserved amino acid sequences among calicivirus capsid proteins.";
 RL Virus Res. 24:211-222(1992).
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; M87481; AAA16217.1; -;
 PIR; A48562; A48562.
 InterPro; IPR004005; Calici coat.
 InterPro; IPR008975; Viral_cap_coat.
 DR PF00915; Calici_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 208 208
 FT CARBOHYD 481 481
 FT CARBOHYD 493 493
 FT CARBOHYD 545 545
 FT CARBOHYD 545 545
 SEQUENCE 702 AA; 77950 MW; E6E5A58523DEE3D7 CRC64;

Query Match 10.4%; Score 300.5; DB 1; Length 702;
 Best Local Similarity 25.2%; Pred. No. 5.8e-15;
 Matches 137; Conservative 66; Mismatches 209; Indels 131; Gaps 25;
 Qy 13 DGTSGAGQLVPEANTAPISMEPVAGATAATAATAGQVNMID-PW-IMNNYVQAPQGETI 70
 Db 154 DPGGADIVTBEQGTVVQQQPVPAQSALTTLAAASTGKTVDCEWTTFFSYHTA-----VNM 209
 Qy 71 SPNNTPGDILFDLQGLPHNPLSLAQMNGWGNMKVLLAGNAFTAGKIISCIPP 130
 Db 210 STTAQKILSPRALSPELNPLRHISLSYTWGSGDVRFTVSGSGVFGKLAALIVPP 269
 Qy 131 GFAAQNTISIAQATMFPFHVIADVRLVLEPIEVPLEVNRVLFHNNNAPTRMLVCMLYTPL- 189
 Db 270 GI--EPVESPTMLQYPHVLFDAQTPEVIFTIPDIRKTLXHSMDTDTTLRLVIMVYNELI 327
 Qy 190 -----RAGSSSGTDPFVIAGRVLTCPSPDFSLVLPVPPNVQKTPSPVNP--LPLNT 241
 Db 328 NPYEQSEPKSCSIT-----VETRPSSDFTFSLKPPG--SLLKHGSIPLDIPRNS 377
 Qy 242 --LNSNRVPSLISKSMVSRDHGMVQVQFNGR-----VTLD----- 274
 Db 378 RHWGNRNMSTIDGFVV-----QPRVFSNRHFPDSTTTGWSTPYYPIPIEVLTKLDRG 432
 Qy 275 QLOQGTTPTSASQL-----CKIRGSVFHANGGYNLTTEL-----DGSPYHAFESPA 321
 Db 433 QYFKVDTTEKSLVPLGDLGWPDTTIPATMTASNGVDTYVAEYRITNNGTHFKGYIMG 492
 Qy 322 PI-----GFPDLGECMDHMEASPTTQNTGDIKQINVKQESAFAP-----HIGTTQA 369

Db	493	NLTKVKGSDNLGET---	QOTSRTLFAVSG-----	NYKQNTINPHTKITSSLVVYDA	543
Qy	370	DGLSDSVSVNT-----	NMTAKLGWV-----	SPVSDGHRGVDVFWIPRYGSTLTLEAAQLA--P	419
Db	544	NNVSAATAKATTTWISTHLLGVLVDSPPV-----	-----	GSDSTKVVRIATLP	587
Qy	420	PIYPPGGEAIVFFWMSDPPIAHGTNGLSVFCTIPQEFVTHFVNEQAPTRGEAALLHY-LD	478		
Db	598	EAFYNG-GNFPVFFNKIQIGH-----	-----	FDRHTKCFNSQVLMTSQKLAENHYTLP	636
Qy	479	PDT	481		
Db	637	PDS	639		
RESULT 5					
Db	POLN MANCV	STANDARD;	PRT;	2208 AA.	
Qy	AC Q69014;				
Db	DT 01-NOV-1997 (Rel. 35, Created)				
Qy	DT 01-NOV-1997 (Rel. 35, Last sequence update)				
Db	DT 28-FEB-2003 (Rel. 41, Last annotation update)				
Qy	DE Genome polypeptide [contains: RNA-directed RNA polymerase				
Db	DE (EC 2.7.7.48); Thiol protease 3C (EC 3.4.22.-); Helicase (2C like				
Qy	DE protein); Coat protein.				
Db	OS Manchester virus (Human enteric calicivirus).				
Qy	OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;				
Db	OC Sapovirus.				
Qy	OC NCBJ_TaxID=82659;				
Db	OC [1]				
Qy	RP SEQUENCE FROM N.A.				
Db	RA MEDLINE=95390791; PubMed=7661689;				
Qy	RA Liu B.L., Clarke I.N., Caul E.O., Lambden P.R.;				
Db	RT "Human enteric caliciviruses have a unique genome structure and are				
Qy	RT distinct from the Norwalk-like viruses.";				
Db	RL Arch. Virol. 140:1345-1356(1995).				
Qy	CC -1- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).				
Db	CC -1- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR				
Qy	CC THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY				
Db	CC SIMILARITY).				
Qy	CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +				
Db	CC (RNA)(N).				
Qy	CC -1- PM: Specific enzymatic cleavages in vivo yield mature proteins.				
Db	CC -1- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.				
Qy	CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.				
Db	CC -----				
Qy	CC This SWISS-PROT entry is copyright. It is produced through a collaboration				
Db	CC between the Swiss Institute of Bioinformatics and the EMBL outstation -				
Qy	CC the European Bioinformatics Institute. There are no restrictions on its				
Db	CC use by non-profit institutions as long as its content is in no way				
Qy	CC modified and this statement is not removed. Usage by and for commercial				
Db	CC entities requires a license agreement (see http://www.isb-sib.ch/announce/				
Qy	CC or send an email to license@isb-sib.ch).				
Db	CC -----				
Qy	CC EMBL; X66560; CAA60262.1; ..				
Db	CC MEROPS; C24.UPW; ..				
Qy	CC InterPro; IPR004005; Calici_coat				
Db	CC InterPro; IPR004004; Calici_pol_hel.				
Qy	CC InterPro; IPR000317; Peptidase_C24.				
Db	CC InterPro; IPR000605; RNA_helicase.				
Qy	CC InterPro; IPR007095; RNA_pol_DS_PS.				
Db	CC InterPro; IPR001205; RNA_pol_P3D.				
Qy	CC InterPro; IPR007094; RNA_pol_PSVir.				
Db	CC InterPro; IPR008975; Viral_cap_coat.				
Qy	CC Pfam; PF00915; Calici_coat; 1.				
Db	CC Pfam; PF03510; Peptidase_C24; 1.				
Qy	CC Pfam; PF00680; RNA_dep_RNA_pol; 1.				
Db	CC Pfam; PF00910; RNA_helicase; 1.				
Qy	CC PRINTS; PR00916; 2GENDOPTASE.				
Db	CC PRINTS; PR00918; CALICIVIRUSNS.				
Qy	CC Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;				
Db	CC Thiol protease; Helicase; ATP-binding; Coat protein.				
Qy	CC Thiol protease; Helicase; ?				
Db	CC CHAIN				

FT	CHAIN	?	?	PROTEASE (P3C).	
FT	CHAIN	?	?	RNA-DIRECTED RNA POLYMERASE.	
FT	CHAIN	?	?	COAT PROTEIN.	
FT	NP BIND	408	415	ATP (POTENTIAL).	
FT	ACT_SITE	1097	1097	PROTEASE (BY SIMILARITY).	
FT	ACT_SITE	1112	1112	PROTEASE (BY SIMILARITY).	
SQ	SEQUENCE	2208 AA;	242736 MW;	3E299D5BA20E45DC CRC64;	
Query Match					
Best Local Similarity 22.9%; Pred.No. 2e-13;					
Matches 138; Conservative 82; Mismatches 244; Indels 138; Gaps 25;					
Qy	14	GTSGAGQLVPEANTAEPISEMPVAGAA-----	TAAATAGQVNMIDPWIMNNVYVQAPQGEFT	69	
Db	1670	GTTG-----	PTTSHVVVANPEQPNGAARLEAVATGA-----	IQSNVPEAIRNCFA	1716
Qy	70	I-----	SPNNTPGDILFDLQGLPHLPPFLSHLAQMYNGVGNMKVLLAGNAFTAGKI	123	
Db	1717	VFRTEFAWDRMPTGTGLGSLSLHENINPYTSLSGMWAGMGSGFEVRLSLISGSGVFAGRI		1776	
Qy	124	IISCIPPGFAAQNISIAQATMPPHVIADVRLIEPIEVLPLEDVNVLFFHNNNA--PTMRLV		182	
Db	1777	IASVDPG--VDPSSIRDPGVLPFAFVDARITEPVSMIPDVRAVDVHRMDGAEPTCSLG		1834	
Qy	183	CMLYTELRASGSSGTDPPFVIAGRVLTCPSPDFSLFLVPP--NVEQTKPFVSNPLN		240	
Db	1835	FWYQPLLPNPFSTTAVSTCWVS--VETKPGGDFDCLLRPPGQOMENGVSPEGLLPRLG		1892	
Qy	241	TLNSRVPSLIKMMVSRDHQWVQ--FONGRVTLDGQLQGTTPTSASOLCKIRGSVFHAN		299	
Db	1893	YSRGNRVGGLVGMILVAEHKQVNRHFNNSVTFGNSVTAQVNPMAA---EIVTQAHST		1948	
Qy	300	GGNGY-----	NLTTEL-DGSPVHAFESPAPI-----	GFPDLGECMDHWEASPTTQ	342
Db	1949	SRHAWLSIGAQNKGFLPFGIPNHPDSCASTVVGAMDTSLGCRPSTGVC-----	GPAIS	2002	
Qy	343	F-NTGDVVKIINVKQESAFAPH-----	LGTIQADGLSDSV-----	N	378
Db	2003	FQNGGDYVE--NDTPSVMFATYDPLTSGTGVAUTNSINPASILALVIRISNNDDPTSGFAND		2060	
Qy	379	TNMTAKLGWSPVSDGH-RGDVDVFWIPRYGSTLTAAQLA-----	PIIYPP	424	
Db	2061	KNVVQMSWEMYGTGNIQIRGVTPMSGTNYTFTSTGANTLVLMQERMLSYDGHQAILYSS		2120	
Qy	425	GFGEAIVFMSDFPIAHGTNGLSVPTCIPOEFVTHFVNEQAPTRGEAALLHYLDPTHRN		484	
Db	2121	QLERTAEYFQNDI-----	VNIPENSMVAF-----	NVETNSA	2151
Qy	485	LGEFKLYPEGEMTCVPNSSGTGPOTLPINGVVFVSVWSRFRYQLKPVGTA--GPACRLGI		542	
Db	2152	SFQIGIRPDGYV-----	TGSGIGNVPLEPETRF-----	QYVGILPLSALSALSGPSGNMGR	2202
Qy	543	RR	544		
Db	2203	AR	2204		
RESULT 6					
COAT_FCVF9	STANDARD;	PRT;	671 AA.		
ID_COAT_FCVF9					
AC_P27406;					
DT_01-AUG-1992 (Rel. 23, Created)					
DT_01-AUG-1992 (Rel. 23, Last sequence update)					
DT_01-JUN-1994 (Rel. 29, Last annotation update)					
DE_Coat protein (Capsid protein).					
GN_CYP76.					
OS_Feline calicivirus (strain F9) (FCV).					
OC_Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;					
OC_Vesivirus.					
OX_NCBJ_TaxID=11981;					
RN_[1]					
RP_SEQUENCE FROM N.A.					
RX_MEDLINE=92410623; PubMed=1529544;					


```

PX MEDLINE=91374597; PubMed=1716692;
RA Neill J.D., Reardon I.M., Heinrichson R.L.;
RT "Nucleotide sequence and expression of the capsid protein gene of
RL feline calicivirus.";
RL J. Virol. 65:5440-5447 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Neill J.D.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; M32819; AAA42925.1; -.
DR EMBL; U13992; AAC13993.1; -.
DR FIR; A40507; VCMWFF.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SSEQUENCE 668 AA; 73550 MW; 9E52312108D4D441 CRC64;
SQ
Query Match 9.3%; Score 269; DB 1; Length 668;
Best Local Similarity 31.0%; Pred. No. 1.3e-12;
Matches 74; Conservative 38; Mismatches 83; Indels 44; Gaps 9;
QY 23 PEANT-AEPTSMPEVAGAAATAAATAGQVNMIDPWIMNNVQAPQGEFTI-----SPNNT 75
DB 133 PEQGTWGVGVIAPFNAQMSAADMATCKSVDSW-----EAFPSHTSVNWSSTST 183
QY 76 PGDILFDLQGLPHNPFSLHQAQVNGVNMKVKVLLAGNAFTAGKIISCIPPGF-AA 134
DB 184 QGKILFQSLGGLNPLNTHLAKLYVANGSVDRFSISGSGVFGKLAIVVPPGIDPV 243
QY 135 QNTSIAQATFPHVIAADVRLVLEIEVPLEDVRNVLPHNNDNAPTMLVCMY-----TPLR 190
DB 244 QSTSMQLQ---YPHVLFDAQVPEVIFSIPLRSTLYHLSMDTDTTSLVIMVYNDLINPYA 300
QY 191 ASGSSGTDPFVIAAGRVLTCPSPDFSLFLVPPNVQKTFPSVPLNLTLSNSRVPS 249
DB 301 NDSNSSGC---IVT--VETKPGDFKPHLLKPPG-----SMLTHGSIPS 339
RESULT 9
COAT_SMSV4 STANDARD; PRT; 703 AA.
ID COAT_SMSV4
AC P36285;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Coat protein (Capsid protein).
OS San Miguel sea lion virus (serotype 4) (SMSV 4).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=36407;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92410750; PubMed=1529644;
RA Neill J.D.;
RT "Nucleotide sequence of the capsid protein gene of two serotypes of

```

```

RT San Miguel sea lion virus: identification of conserved and non-
RT conserved amino acid sequences among calicivirus capsid proteins.";
RL Virus Res. 24:211-222 (1992).
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; M87482; AAA16220.1; -.
DR FIR; C48562; C48562.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 482 482 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SSEQUENCE 703 AA; 77721 MW; C5DAD8223B261073 CRC64;
SQ
Query Match 9.1%; Score 264; DB 1; Length 703;
Best Local Similarity 27.4%; Pred. No. 3.3e-12;
Matches 104; Conservative 53; Mismatches 181; Indels 42; Gaps 17;
QY 13 DGTSGAGQLVPEANTAEPISEMEPVAGAAATAAATAGQVNMIDPWIMNNVQAPQGEFTISP 72
DB 154 DGPSSAEIVTEEGTVVQQPAPAPALATLATATAGTCKSVGEQEMWTF--SYHTSINWST 211
QY 73 NNTPGDILFDLQGLPHNPFSLHQAQVNGVNMKVKVLLAGNAFTAGKIISCIPPGF 132
DB 212 VESQGLTLYSQALNPSINPVLHIAKLYSTWSGGIDVREFTVSGSGVFGKLAALLVPPGV 271
QY 133 -AAQNTSIAQATFPHVIAADVRLVLEIEVPLEDVRNVLPHNNDNAPTMLVCMYTPPLA 191
DB 272 EPIESVSMQLQ---YPHVLFDAQVPEVIFSIPLRSTLYHLSMDTDTTSLVIMVYNDLINPY 324
QY 192 SGSSSGTDPFVIAAGRVLTCPSPDFSLFLVPPNVQKTFPSVPLNLTLSNSRV 247
DB 325 NVGENKTTCSI---TVETKPSADFTFALLKPPG--SLIKHSIPSDLIIPNSAHMGNRW 379
QY 248 PSLLIKSMVSRDHQVQVQNGR--VTLDGQLQG--TFTSASQLCKIRGSHFANGNG-- 303
DB 380 WSTISGFSV-----QPRVFSQNRHDFDFTTTCGWSSTPYVPIEIKIQGV-----GSNNKW 430
QY 304 YNLTELDGSPYHAFESAPIGFDPDLGECWHMEASPTT-QFNTGDVVK--QINVKQESAF 360
DB 431 FHVITD-----KALVGPIDGWDPTTIPD---ETKATNGNFSYGESVRAGSTTIKPNENS 483
QY 361 APHLGTIQADGLSDVSVNTN 380
DB 484 THFKGYTCGLSTVEIPEN 503
RESULT 10
POLG_POL2W STANDARD; PRT; 2205 AA.
ID POLG_POL2W
AC P23069;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (PIA); Coat protein VP2
DE (PIB); Coat protein VP3 (PIC); Coat protein VP1 (PID); Picornain 2A
DE (EC 3.4.22.29) (P2A); Core protein P2B; Core protein P2C; Core protein
DE P3A; Genome-linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28)
DE (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].
DE Poliovirus type 2 (strain W-2).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC

```

Enterovirus.
 NCBI_TaxID=12085;
 [1]
 SEQUENCE FROM N.A.
 MEDLINE=90155230; PubMed=2154539;
 Peavey D.C., Oh C.K., Cunningham L.L., Calenoff M., Jubelt B.;
 "Localization of genomic regions specific for the attenuated, mouse-
 adapted poliovirus type 2 strain W-2.";
 J. Gen. Virol. 71:43-52(1990).
 CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
 O/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
 CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln|-Gly bond in the
 poliovirus polyprotein. In other picornavirus reactions Glu may be
 substituted for Gln, and Ser or Thr for Gly.
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 (RNA) (N).
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 each of which is composed of one copy each of proteins VP1, VP2,
 VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; D06625; BAA00516.1; ALT_SEQ.
 DR PIR; A34032; GNNY2W.
 DR HSSP; P03299; 1POV.
 DR MEROPS; C03.001; -.
 DR MEROPS; C03.020; -.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR001199; Pept_3C_picorn.
 DR InterPro; IPR000081; Peptidase_C3.
 DR InterPro; IPR003138; Pico_P1A.
 DR InterPro; IPR002527; Pico_P2B.
 DR InterPro; IPR001676; RNV.
 DR InterPro; IPR000605; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_BSvir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00548; Cys-protease-3C; 1.
 DR Pfam; PF02226; Pico_P1A; 1.
 DR Pfam; PF00947; Pico_P2A; 1.
 DR Pfam; PF01552; Pico_P2B; 1.
 DR Pfam; PF00073; rhv_3.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICIVIRUSN.
 DR ProDom; PD001125; Cys_protease_3C; 1.
 DR ProDom; PD001306; Pico_P2A; 1.
 DR ProDom; PD001274; Pico_P2B; 1.
 DR SMART; SM00382; AAA; 1.
 KW Polyprotein; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
 KW Lipoprotein.
 FT CHAIN 2 69 COAT PROTEIN VP4.
 FT CHAIN 70 340 COAT PROTEIN VP2.
 FT CHAIN 341 578 COAT PROTEIN VP3.
 FT CHAIN 579 879 COAT PROTEIN VP1.
 FT CHAIN 1028 1028 PROTEASE 2A.
 FT CHAIN 1029 1125 CORE PROTEIN 2B.
 FT CHAIN 1126 1454 CORE PROTEIN 2C.
 FT CHAIN 1455 1541 CORE PROTEIN 3A.
 FT CHAIN 1542 1563 GENOME-LINKED PROTEIN VPg.
 FT CHAIN 1564 1746 PICORNAIN 3C.

FT	CHAIN	1747	2205	RNA-DIRECTED RNA POLYMERASE 3D.
FT	LIPID	2	2	N-myristoyl glycine (in host) (By similarity).
FT	ACT_SITE	1710	1710	PROTEASE (POTENTIAL).
FT	ACT_SITE	1724	1724	PROTEASE (POTENTIAL).
FT	SEQUENCE	2205	2205	AA; 245701 MW; 2422AB039E0254AD CRC64;
Qy	Query Match	5.1%	Score 146.5; DB 1; Length 2205;	
Qy	Best Local Similarity	19.2%	Pred. No. 0.011;	
Qy	Matches	109;	Conservative 71; Mismatches 108; Indels 199; Gaps 26;	
Qy	22	VPEANTAEPTSMFVAGAAATAAGQVNMIDPMNMYVQAPQGEFTISN-----NTP 76		
Qy	365	IPEFDVTPPID---IPGEVRNMMELAEIDTMIPLNLTISQRKNTWDMVRVELNDAHSPTP 421		
Qy	77	GDILFDLQGLHLPFLSH-----LAQMYNGWGNMKVKVLLAGNAFTAGKIIISCIIPP 131		
Qy	422	---ILCLSLSPASDPRLAHTMLGELILNYTHWAGSLKFTFLFCGSMATGKLLSYAPPG 478		
Qy	132	FAAQNISIAQNTMPHVIADVRVLEPIEVLPLEVVRNVLFNH--NDNAP-----TMR 180		
Qy	479	AKAPE-SRKEAMLGTHVINDIGLQSSCTMVVPMISNTTYRQTINDSFTEGGYISMFYQTR 537		
Qy	181	LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFL-----VDPN----- 224		
Qy	538	VVPELSTPRKMD-----ILGFVSAC--NDFSRLRLDRTTHISQEVNPOGLDLIE 585		
Qy	225	--VEOKTKPFSVPLPLNTLSNR-----VPS-----LIKSMVYSRDHGQ 262		
Qy	586	GVVEGVTRNALTPLTPVNNLPDTRSSGPAHSKETPALTAVETGATNPLVPSDTVQTRH-- 643		
Qy	263	MQV-----FQNGRVTLGQLQGTPT--SASQLCKIRGSVF----- 296		
Qy	644	VIQKTRSESTVESFFARGACVALLIENDVNDAPTRASKLPFSVKITTKDTVOLARKLEFF 703		
Qy	297	-----HANGGNGVNLTELDGSPHAFESPAPIGFPDLG-----ECWHEA 337		
Qy	704	TYSRFDMFTFVTSNTYTDANNGHALNQVQIMY-----IPGAPIPKRNDTWQTSS 757		
Qy	338	SPTTQFNTGDIKQINVKQBSAPAPHLGTTQA-----DGLSDSVSVNTNMIAKLGWSPVS 392		
Qy	758	NPSVFYTYGAPPARISV-----PYGVIANAVSHFYDGFVKV----- 793		
Qy	393	DGHRGDVDPWVIRYRSTLTAQAAPPYPPGGEAIVFFMSDPPIAHGNTGLSVPTCTI 452		
Qy	794	-----PLAQQAQSTEGDSL-----YGAA-----SLNDF-----GSLAV----- 820		
Qy	453	QGEVTHFVNEQAPTRGEAALLHYLDP 479		
Qy	821	-----RVVNDHNPTKLTISKIRVYMKP 841		

RESULT 11

POLG_POL2L
 ID POLG_POL2L STANDARD; PRT; 2207 AA.

AC P06210;

DT 01-JAN-1988 (Rel. 06, Created)

DT 01-JAN-1988 (Rel. 06, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Picornain 2A (EC 3.4.22.29) (P2A); Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VPg (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].

OS Poliovirus type 2 (strain Lansing).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Enterovirus.

OX NCBI_TaxID=12084;

[1]

RP SEQUENCE FROM N.A.

RX MEDLINE=86115399; PubMed=3003384;

RA la Monica N., Meriam C., Racaniello V.R.;

RT "Mapping of sequences required for mouse neurovirulence of poliovirus

RT J. Virol. 57:515-525(1986).
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC O/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/submit/>
CC or send an email to license@ebi.ac.uk).
CC
CC -----
CC EMBL; M12197; AAA46912.1; -.
CC PIR; A29507; GNNY5P.
CC HSSP; P03299; IPOV.
CC MEROPS; C03.001; -.
CC MEROPS; C03.020; -.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR001099; Pept_3C_picorn.
CC InterPro; IPR000081; Peptidase_C3.
CC InterPro; IPR003138; Pico_P1A.
CC InterPro; IPR002527; Pico_P2B.
CC InterPro; IPR001676; Rhv.
CC InterPro; IPR006005; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_PSVir.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00548; Cys-protease_3C; 1.
CC Pfam; PF02226; Pico_P1A; 1.
CC Pfam; PF00947; Pico_P2A; 1.
CC Pfam; PF01552; Pico_P2B; 1.
CC Pfam; PF00073; rhv; 3.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICIVIRUSNS.
CC ProDom; PD001125; Cys_protease_3C; 1.
CC ProDom; PD001306; Pico_P2A; 1.
CC ProDom; PD001274; Pico_P2B; 1.
CC SMART; SM00382; AAA; 1.
CC Polyprotein; Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
CC Lipoprotein.
KW CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 340 COAT PROTEIN VP2.
FT CHAIN 341 578 COAT PROTEIN VP3.
FT CHAIN 579 879 COAT PROTEIN VP1.
FT CHAIN 880 1028 PROTEASE 2A.
FT CHAIN 1029 1125 CORE PROTEIN 2B.
FT CHAIN 1126 1454 CORE PROTEIN 2C.
FT CHAIN 1455 1541 CORE PROTEIN 3A.
FT CHAIN 1542 1563 GENOME-LINKED PROTEIN VP6.
FT CHAIN 1564 1746 PICORNAIN 3C.
FT CHAIN 1747 2207 RNA-DIRECTED RNA POLYMERASE 3D.
FT LIPID 2 2 N-myristoyl glycine (in host).
FT ACT_SITE 1710 1710 PROTEASE (POTENTIAL).
FT ACT_SITE 1724 1724 PROTEASE (POTENTIAL).
SQ SEQUENCE 2207 AA; 245829 MW; 2B1E2070B7D44F99 CRC64;
Query Match 5.0%; Score 145.5; DB 1; Length 2207;

Best Local Similarity 18.8%; Pred. No. 0.013;
Matches 106; Conservative 72; Mismatches 193; Indels 193; Gaps 24;
QY 22 VPEANTAPISMEPVAGAAATAAATAGQVNWIDPMNMNYVQAPOGETISP-----NTP 76
DB 365 IPEFDTVPID-----IPGEVRNMBELADITMPLNLTNQRKNTWDMYRVELNDAHSDTP 421
QY 77 GDILFDLQGLHLPFLSH-----LAQWYNGWGNMKVKVLLAGNAFTAGKIIISCIPIPP 131
DB 422 ---ILCLSLSPASDPLAHTMLGELNYTHWAGSLKFTFLFCGSMATGKLLVSAPP 478
QY 132 FAAQNIASIAQTMPEPHVIADVRVLEPIEVPLEDVRLVFNH-----NDNAP-----TMR 180
DB 479 AEAPK-SRKEAMLGTHVWDIGLQSSCTMVVFWISNTTYRQTINDSFTEGGYISMFYQTR 537
QY 181 LVCMLYPLRASGSSGTDFFVIAGRVLTCPSPDPSPLF-----LVP 222
DB 538 VVVPLSTPRKMD-----ILGFVSAC-----NDFSVELLRDTHISOEAMPQGLDLIE 585
QY 223 PNVEQKTKPFSVNPPLNLTLSNR-----VPS-----LIKSMVSRDHGQ 262
DB 586 GVVEGVTRNALTPLTPANNLPDTQSSGPAHSKETALTAVETGATNPLVPSDTVQTRH-- 643
QY 263 MVQ-----FQGRVTLQGLQGTPTS-ASOLCKIRGSVF----- 296
DB 644 VIQKTRSESTVESPFARGACVAIIEVNDNDAPTCKASKLFSWKITYKDTVQLRRKLEFF 703
QY 297 -----HANGGNGYNLTDLGSPYHAFESPAPIGPFDIGECDDHWEASPT 340
DB 704 TYSRPDMETFFVTSNYTANNHALNQYIWMIPPGAPIGKWNKY-----TWOTSSNPS 760
QY 341 TQFTGDEVKQINVKQESAPAPHLGTIOA-----DGLSDVSNTNMIAGLGVSPVSDGH 395
DB 761 VFTYGAPPARISV-----PVVGIANAYSHFYDGFVKV----- 793
QY 396 RGVDPWVPIRYGSTLTEAAQLAPPIYPFGGEAIVFFMSPDFIAHGTNGLSVPCPTPOE 455
DB 794 -----PLAQASTEGDSL-----YGAA-----SLNDF-----GSLAV----- 820
QY 456 FVTHFVNEQAPTRGEAALLHYLDP 479
DB 821 ---RVVNDHNPTKLTISKIRVYMKP 841
RESULT 12
POLG_POL3L STANDARD; PRT; 2206 AA.
ID POL3302; Q84783; Q84784; Q84785; Q84786; Q84787; Q84788; Q84789;
AC P03302; Q84783; Q84784; Q84785; Q84786; Q84787; Q84788; Q84789;
AC Q84790; Q98592; Q98593; Q98594;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C, P3A; Genome-linked protein VP6; Picornain 3C
DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D
DE (EC 2.7.7.48)].
OS Poliovirus type 3 (strains P3/Leon/37 and P3/Leon 12A[1]B).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12088;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P3/Leon/37;
RX MEDLINE=84170338; PubMed=6324200;
RA Stanway G., Hughes P.J., Mountford R.C., Reeve P., Minor P.D.,
RA Schild G.C., Almond J.W.;
RT "Comparison of the complete nucleotide sequences of the genomes of
RT the neurovirulent poliovirus P3/Leon/37 and its attenuated Sabin
RT vaccine derivative P3/Leon 12alb."
RL Proc. Natl. Acad. Sci. U.S.A. 81:1539-1543 (1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=P3/Leon 12A[1]B;

RX MEDLINE=83299239; PubMed=6310508;
RA Stanway G., Cann A.J., Hauptmann R., Hughes P.J., Clarke L.D.,
RA Mountford R.C., Minor P.D., Schild G.C., Almond J.W.;
RT "The nucleotide sequence of poliovirus type 3 Leon 12 alb: comparison
RT with poliovirus type 1";
RL Nucleic Acids Res. 11:5629-5643(1983).
RN (3)
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 1-878.
RN MEDLINE=95120467; PubMed=7820548;
RX Grant R.A., Hiremath C.N., Filman D.J., Syed R., Andries K.,
RA Hogle J.M.;
RT Structures of poliovirus complexes with anti-viral drugs:
RT implications for viral stability and drug design.";
RL Curr. Biol. 4:784-797(1994).
RN (4)
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1-878.
RA Hiremath C.N., Grant R.A., Filman D.J., Hogle J.M.;
RT "Binding of the antiviral drug wins1711 to the Sabin strain of type-3
RT poliovirus - structural comparison with drug-binding in rhinovirus-
RT 14.";
RL Acta Crystallogr. D 51:473-489(1995).
CC -|- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC O/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -|- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -|- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -|- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -|- MISCELLANEOUS: THE SEQUENCE OF STRAIN SABIN VACCINE P3/LEON/37 IS
CC SHOWN.
CC -|- MISCELLANEOUS: THE STRAIN SABIN VACCINE P3/LEON/37 IS THE
CC PROGENITOR OF THE STRAIN SABIN VACCINE P3/LEON 12A[1]B.
CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K01392; AAA46914.1; -;
DR EMBL; X00925; CAA25444.1; -;
DR PDB; 1P1V; 03-JUN-95.
DR PDB; 1PVC; 15-SEP-95.
DR PDB; 1VBA; 11-JUL-96.
DR PDB; 1VBB; 11-JUL-96.
DR PDB; 1VBC; 11-JUL-96.
DR PDB; 1VBE; 11-JUL-96.
DR MEROPS; C03.001; -;
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000199; Pept_3C_picorn.
DR InterPro; IPR000081; Peptidase_C3.
DR InterPro; IPR003138; Pico_P1A.
DR InterPro; IPR002527; Pico_P2B.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_P5.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_P5vir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00548; Cys-protease-3C; 1.
DR Pfam; PF02226; Pico_P1A; 1.
DR Pfam; PF00947; Pico_P2A; 1.
DR Pfam; PF01552; Pico_P2B; 1.
DR Pfam; PF00073; rhv; 3.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PD00918; CALICIVIRUSNS.
DR ProDom; PD001125; Cys_protease_3C; 1.
DR ProDom; PD001306; Pico_P2A; 1.
DR ProDom; PD001274; Pico_P2B; 1.
DR Polyprotein; Coat protein; Core protein; Transferase;
DR RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
DR 3D-structure; Lipoprotein.
FT CHAIN 2 69
FT CHAIN 70 340
FT CHAIN 341 578
FT CHAIN 579 878
FT CHAIN 879 1027
FT CHAIN 1028 1124
FT CHAIN 1125 1453
FT CHAIN 1454 1540
FT CHAIN 1541 1562
FT CHAIN 1563 1745
FT CHAIN 1746 2206
FT LIPID 2 2
FT ACT SITE 1709 1709
FT ACT_SITE 1723 1723
FT VARIANT 431 431
FT VARIANT 864 864
FT VARIANT 908 908
FT STRAND 4 8
FT STRAND 25 29
FT STRAND 36 38
FT TURN 50 50
FT HELIX 51 54
FT STRAND 57 57
FT STRAND 63 64
FT TURN 80 81
FT STRAND 83 87
FT TURN 88 89
FT STRAND 90 94
FT STRAND 101 102
FT STRAND 103 105
FT TURN 113 115
FT STRAND 123 123
FT HELIX 126 128
FT TURN 129 129
FT STRAND 133 134
FT STRAND 138 141
FT TURN 142 143
FT STRAND 147 151
FT TURN 152 152
FT HELIX 153 155
FT TURN 156 157
FT HELIX 159 167
FT STRAND 168 180
FT TURN 185 186
FT STRAND 187 197
FT TURN 198 198
FT STRAND 203 203
FT TURN 208 209
FT HELIX 213 216
FT HELIX 219 221
FT STRAND 223 224
FT STRAND 226 226
FT TURN 235 236
FT STRAND 242 242
FT STRAND 245 245
FT HELIX 246 248
FT TURN 249 252
FT HELIX 255 260
FT STRAND 263 267
FT TURN 268 270
FT STRAND 273 278
FT STRAND 287 287
FT TURN 289 291
FT STRAND 292 292

FT	STRAND	295	307
FT	TURN	308	309
FT	STRAND	310	310
FT	STRAND	314	331
FT	STRAND	337	337
FT	TURN	348	351
FT	TURN	355	356
FT	STRAND	363	363
FT	TURN	366	367
FT	STRAND	379	380
FT	STRAND	382	382
FT	TURN	383	383
FT	HELIX	384	387
FT	TURN	388	388
FT	STRAND	391	392
FT	STRAND	397	397
FT	TURN	399	403
FT	HELIX	405	408
FT	STRAND	410	413
FT	TURN	414	415
FT	TURN	418	419
FT	STRAND	423	426
FT	TURN	429	431
FT	TURN	433	437
FT	HELIX	439	444
FT	TURN	445	446
FT	STRAND	447	451

Query Match 5.0%; Score 145; DB 1; Length 2206;
Best Local Similarity 20.5%; Pred. No. 0.015;
Matches 98; Mismatches 194; Indels 128; Gaps 20;

Qy	15	TSGAGQLVPEANTAEPISMEPVAGATAAATAGOV-NM-----IDPWIMNNVQAPOG--	66
Db	347	TPGSNQYITSDNHQSFCAL-PEFDVTPIPIDPGEVKNMELAEIDTMIPLNESTKRTM	405
Qy	67	---EFTISPNNTPGDILFDLQLGPHLNPFSLH-----LAQMNGWGVGNMKVLLAGNAF	118
Db	406	DMRYVTLSDSADLSQPILCLSLSPASDPRLSHTMLGEVLNYTHWAGSLKFTFLFCGSM	465
Qy	119	TAGKIISCIPGPAAQNISIAQTAFMHVIADVLEPIEVPLEDVNVLFH--NNDNA	176
Db	466	ATGLKLVAYPEG-AQPPTSRKEAMLTGHVMDLGLQSSCTMVVPWISNVIVRQTQDSF	524
Qy	177	PMRLVCMLY-----TPLRASSGSSGTDPPFIAGRVLTCPSDFSFLEL-----V	221
Db	525	TGGYISMFYQRIVVPLSTPKSMS-----MLGFVSAC--NDFSRLLRDTTHISQSAL	576
Qy	222	PRNVE-----QKTQFPFVP---NLP-----LATLSNSRVPSLIKNMM	255
Db	577	PQIEDLISEVAQGALTLSLPKQDSLPTKASGPAHSKEVPALTAVETGATNPAPSOT	636
Qy	256	VSRDHQMWQ-----FONGRVTLDGLOQTTPTSASQLCKIRGSVFH-----	297
Db	637	VQTRH--VVQRRSRSESTIESFARGACVAIIIEVDNEQFTTRAQKLFAMWRITYKDTVOL	694
Qy	298	-----ANGNGYNILTLDGSPYHAFESPAGIPDLGECDM	333
Db	695	RRKLEPFTYSRFDMFEFTFVVTAFTNANHGHALNQYIQIMIYPFGAPTKSMDDY---TW	751
Qy	334	HMEASTPTQNTGDVIKQINVKESAFAPHLGTIOA-----DGLSDSVSNVTNNIAKLG	386
Db	752	QTSNPSIFITYGAAPARISV-----FVGLNAYSHEYDFGFAKVPLKTDANDQIG	802

RESULT 13
 POLG_HE701
 ID_POLG_HE701 STANDARD; PRT; 2194 AA.
 AC_P32537;
 DT_01-OCT-1993 (Rel. 27, Created)
 DT_01-NOV-1995 (Rel. 32, Last sequence update)
 DE_10-OCT-2003 (Rel. 42, Last annotation update)
 DE_Genome polypeptide [Contains: Coat protein VP4 (P1A); Coat protein VP2]

each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.

-!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.

-!- MISCELLANEOUS: THIS VIRUS IS A LIVE VACCINE STRAIN DERIVED FROM THE MAHONEY STRAIN BY SPONTANEOUS MUTATIONS DURING THE ATTENUATION PROCESS.

-!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

EMBL; V01150; CAA24465.1; -.

MEROPS; C03.001; -.

MEROPS; C03.020; -.

InterPro; IPR004004; Calici_pol_hel.

InterPro; IPR009003; Cys_ser_trypsin.

InterPro; IPR000199; Pept_3C_picorn.

InterPro; IPR000081; Peptidase_C3.

InterPro; IPR003138; Pico_P1A.

InterPro; IPR002527; Pico_P2B.

InterPro; IPR001676; Rhv.

InterPro; IPR000605; RNA_helicase.

InterPro; IPR007095; RNA_pol_DS_PS.

InterPro; IPR001205; RNA_pol_P3D.

InterPro; IPR007094; RNA_pol_P5vir.

InterPro; IPR008975; Viral_cap_coat.

Pfam; PF00548; Cys-protease_3C; 1.

Pfam; PF02226; Pico_P1A; 1.

Pfam; PF00947; Pico_P2A; 1.

Pfam; PF01552; Pico_P2B; 1.

Pfam; PF00073; rhv_3.

Pfam; PF00680; RNA_dep_RNA_pol; 1.

Pfam; PF00910; RNA_helicase; 1.

PRINTS; PR00918; CALICIVIRUS.

ProDom; PD001125; Cys_protease_3C; 1.

ProDom; PD001306; Pico_P2A; 1.

ProDom; PD001274; Pico_P2B; 1.

Polyprotein; Coat protein; Core protein; Transferase; RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate; Lipoprotein.

CHAIN 2 69 COAT PROTEIN VP4.

CHAIN 70 341 COAT PROTEIN VP2.

CHAIN 342 579 COAT PROTEIN VP3.

CHAIN 580 881 COAT PROTEIN VP1.

CHAIN 882 1030 CORE PROTEIN P2A.

CHAIN 1031 1127 CORE PROTEIN P2B.

CHAIN 1128 1456 CORE PROTEIN P2C.

CHAIN 1457 1543 CORE PROTEIN P3A.

CHAIN 1544 1565 GENOME-LINKED PROTEIN VPG.

CHAIN 1566 1747 PICORNAIN 3C.

CHAIN 1748 2209 RNA-DIRECTED RNA POLYMERASE P3D.

LIPID 2 1712 N-myristoyl glycine (in host).

ACT SITE 1712 1712 PROTEASE (POTENTIAL).

ACT SITE 1726 1726 PROTEASE (POTENTIAL).

SEQUENCE 2209 AA; 246576 MW; 9EC1EF4D174A28A4 CRC64;

Query Match 4.7%; Score 137.5; DB 1; Length 2209;

Best Local Similarity 19.3%; Pred. No. 0.054;

Matches 108; Conservative 73; Mismatches 193; Indels 187; Gaps 25;

QY 22 VPEANTAEPISEPVAGATAAATAGVNMIDPMNMNYVOAPOGEFTI---SPNTPGD 78

DB 365 IPEDVTPPID---IPGEVKNVMEIAEDITMPLNLENTKNTWDMYRVRLSDSANLSGP 421

QY 79 ILFDLQGLPHLNPFLSH-----LAQMYNGVGNMKVILLAGNAFTAGKIIISCIPGFA 133

DB 422 ILC-LSLSPAADPRLSHTMLGEVLNYTHWAGSLKFTFLFCGSMWATGKLVAAPG-A 479

QY 134 AQTLSIAQATWPHVIAVRVLEPIEVLVDVNRVLFH--NNDNAPTMRVLCMLY----- 186

DB 480 QPPTSRKAMLGTHVINDLGLQSSCTMVVPHISNVTYRQTQDSFTGGYISMFYQTRIV 539

QY 187 TPLRASGSSGTDPPFVIAGRVLTCPSPDFSLFL-----VPPNVE-----QKT 229

DB 540 VPL-----STPKAMDMLGFVSAC--NDFSRLRLDTHISQAAMPQGVDDLITEVAQNA 591

QY 230 KFSVSP---NLP-----LNTLS-----NSRVPS 249

DB 592 LALSLEKPSQLPDTKASGPAHSKEVPTLTAVETGATNPLVPSDTVQTRHVHQRSSES 651

QY 250 LKSM-----MVSRDHQ---MVQ--FQNGRVT-----LDQGLQGTFTTASQLCK 290

DB 652 TIESFARGACVAIIIEVDNEQPATNVQKLFATWRITYKDTVQLRKLKFFTVSRFDMFT 711

QY 291 IGRSVFHANGNGYNLTEDLGSFYHAFESPAIGFDPDLGECWHEASPTTQFNTGDIK 350

DB 712 FVVTANFTNSNGHALNOVYQIYIPPGAPTCKSDWDY---TWQTSNPSIFTYTGAAPA 768

QY 351 QINVKQESAFAPHLGTIOA-----DGLSDVSVNTNNIAKLG 386

DB 769 RISV-----PVGLANAYSHFYDGFVPLKSDANDQVG 802

RESULT 15

POLG_POLLS STANDARD; PRT; 2209 AA.

ID POLG_POLLS Q84881; Q84882; Q84883; Q84884; Q84885; Q84886; Q84887;

AC Q84888; Q84889; Q84890;

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Genome polyprotein (Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Picornain 2A (EC 3.4.22.29) (P2A); Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VPG (P3B); Picornain 3C (EC 3.4.22.28) (P3D)).

DE Poliovirus type 1 (strain Sabin).

OS Poliovirus type 1 (strain Sabin).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.

OC NCBI_TaxID=12082;

OX [1]

RN SEQUENCE FROM N.A.

RA MEDLINE=83299876; PubMed=6310545;

RA Nomoto A., Omata T., Toyoda H., Kuge S., Horie H., Kataoka Y.,

RA Genba Y., Nakano Y., Imura N.;

RT "Complete nucleotide sequence of the attenuated poliovirus Sabin 1 strain genome.";

RT Proc. Natl. Acad. Sci. U.S.A. 79:5793-5797(1982).

RL FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.

CC CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the poliovirus polyprotein. In other picornavirus reactions Gln may be substituted for Gln, and Ser or Thr for Gly.

CC CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA} (N).

CC SUBUNIT: The virus capsid is composed of 60 icosahedral units,

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:35:27 ; Search time 31.7307 Seconds
(without alignments)
5419.275 Million cell updates/sec

Title: US-09-926-799-1
Perfect score: 2896
Sequence: 1 MMASKDAPTNMDGTSGAQ.....YOLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL_25:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2873	99.2	545	12	Q8VA02 human calic
2	2846	98.3	545	12	Q918C8 norwalk-lik
3	2846	98.3	545	12	Q9W182 norwalk-lik
4	2840	98.1	545	12	Q918C2 norwalk-lik
5	2798	96.6	545	12	Q91H12 human calic
6	2744	94.8	545	12	Q91V45 human calic
7	2605	90.0	544	12	Q66418 desert shie
8	2139	73.9	543	12	Q8BCA3 human calic
9	2124	73.3	539	12	Q91V48 human calic
10	2017.5	69.7	543	12	Q8JXJ0 norwalk-lik
11	1979	68.3	544	12	Q9DU46 chiba virus
12	1977.5	68.3	546	12	Q8JW44 norwalk-lik
13	1975	68.2	544	12	Q91I85 human calic
14	1974	68.2	544	12	Q9QTE7 chiba virus
15	1972.5	68.1	540	12	Q9YS14 norwalk vir
16	1968	68.0	544	12	Q918C5 norwalk-lik

17	1964.5	67.8	546	12	Q9IV47	Q9IV47 human calic
18	1956.5	67.6	543	12	Q9IV43	Q9IV43 human calic
19	1956.5	67.6	544	12	Q8VA27	Q8VA27 human calic
20	1955.5	67.5	540	12	Q8BCA5	Q8BCA5 human calic
21	1954	67.5	544	12	Q918D1	Q918d1 norwalk-lik
22	1952	67.4	544	12	Q918J2	Q918j2 human calic
23	1946.5	67.2	543	12	Q918B9	Q918b9 norwalk-lik
24	1946	67.2	530	12	Q83884	Q83884 norwalk vir
25	1942	67.1	544	12	Q91V36	Q91V36 human calic
26	1941	67.0	544	12	Q91V41	Q91V41 human calic
27	1938	66.9	530	12	Q9QT39	Q9qt39 norwalk-lik
28	1930.5	66.7	531	12	Q8V7J5	Q8v7j5 norwalk-lik
29	1908.5	65.9	517	12	Q9JH41	Q9jH41 norwalk-lik
30	1904	65.7	539	12	Q8JVV5	Q8jVv5 norwalk-lik
31	1900	65.6	530	12	Q83876	Q83876 norwalk vir
32	1886	65.1	541	12	Q91V42	Q91V42 human calic
33	1288	44.5	522	12	Q8B558	Q8b558 bovine ente
34	1285	44.4	522	12	Q8B4Z5	Q8b4z5 norwalk-lik
35	1276	44.1	522	12	Q8B4Z2	Q8b4z2 norwalk-lik
36	1274	44.0	522	12	Q8V629	Q8v629 norwalk-lik
37	1269	43.8	522	12	Q8B4Y9	Q8b4y9 norwalk-lik
38	1268	43.8	522	12	Q8V628	Q8v628 norwalk-lik
39	1261	43.5	520	12	Q7T8L1	Q7t8l1 bovine ente
40	1260.5	43.5	519	12	Q9YQ22	Q9yq22 bovine cali
41	1256	43.4	520	12	Q7TBK7	Q7tbk7 bovine ente
42	1228	42.4	540	12	Q917X1	Q917x1 norwalk-lik
43	1223	42.2	542	12	Q66293	Q66293 calicivirus
44	1222	42.2	542	12	Q913B6	Q913b6 human calic
45	1219	42.1	540	12	Q91V50	Q91V50 human calic

ALIGNMENTS

RESULT 1

Q8VA02	PRELIMINARY;	PRT;	545 AA.
ID	Q8VA02		
AC	Q8VA02;		
DT	01-MAR-2002 (TremBLrel. 20, Created)		
DT	01-MAR-2002 (TremBLrel. 20, Last sequence update)		
DT	01-OCT-2003 (TremBLrel. 25, Last annotation update)		
DE	Capsid protein.		
OS	Human calicivirus NLV/Potsdam 196/2000/DE.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;		
OC	Norovirus.		
OX	NCBI_TaxID=150135;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HU/NLV/Potsdam 196/2000/DE;		
RA	Kunkel U, Schlegel P, Eg		
RT	Molecular epidemiology of outbreaks of gastroenteritis associated		
RT	with Norwalk-like viruses in Germany."		
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF439287; AAL32455.1; -		
DR	InterPro; IPR004005; Calici.coat.		
DR	InterPro; IPR008975; Viral_cap.coat.		
DR	Pfam; PF00915; Calici_coat_1		
SQ	SEQUENCE 545 AA; 58601 MW; 7D47D4D9F3521D89 CRC64;		

Query Match 99.2%; Score 2873; DB 12; Length 545;
Best Local Similarity 99.1%; Pred. No. 2.4e-230;
Matches 540; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy	1	MMASKDAPTNMDGTSGAGQLVPEANTAEPTSMSPVAGATAAATAGOVNMDPWIMNY	60
Db	1	MMASKDAPTNMDGTSGAGQLVPEANTAEPTSMSPVAGATAAATAGOVNMDPWIMNY	60
Qy	61	VQAPQGEFTTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVKVLLAGNAFTA	120
Db	61	VQAPQGEFTTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVKVLLAGNAFTA	120
Qy	121	GKIIISCIIPGFAQNTISIAQATPFIADRVLEPIEVPLEDVRNVLFNNNNVPTMR	180


```
Query Match      98.3%; Score 2846; DB 12; Length 545;
Best Local Similarity 98.3%; Pred. No. 4.3e-228;
Matches 536; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 MMASKDPTNMDTSGAGQLVPEANTAEPISEMPVAGAAATAAGVNNIDPWIMNY 60
DB 1 MMASKDPTNMDTSGAGQLVPEANTAEPISEMPVAGAAATAAGVNNIDPWIMNY 60
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHNPFSLHQAQYNGWGNMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHNPFSLHQAQYNGWGNMKVKVLLAGNAFTA 120
QY 121 GKIIISCIIPGFAAQNISIAQATMPHVIADRVLEPIEVEPLEDVNRNVLFRNNNDNPTMR 180
DB 121 GKIIISCIIPGFAAQNISIAQATMPHVIADRVLEPIEVEPLEDVNRNVLFRNNNDNPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPPVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSVPNLPLN 240
DB 181 LVCMLYTPLRASGSSGTDPPVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSVPNLPLN 240
QY 241 TLSNRVPSLTKSMVSRDHQGMVQFNGRVTLTGQGLTPTTSASQLCKIRGSVFHANG 300
DB 241 TLSNRVPSLTKSMVSRDHQGMVQFNGRVTLTGQGLTPTTSASQLCKIRGSVFHANG 300
QY 301 GNGYNLTDLGSPYHAFESPAPIGFPDLGECDMHMEASPTTQFNTGDIKQINVKQESAF 360
DB 301 GNGYNLTDLGSPYHAFESPAPIGFPDLGECDMHMEASPTTQFNTGDIKQINVKQESAF 360
QY 361 APHLGTIOADGLSDVSVNTNMIKLGWVSPVSDGHRGVDVDPWIPRYGSTITEAAQLAPP 420
DB 361 APHLGTIOADGLSDVSVNTNMIKLGWVSPVSDGHRGVDVDPWIPRYGSTITEAAQLAPP 420
QY 421 IYPGFGGAIIVFFMSDFPIAHGTNGLSVPCCTIPOEFVTHFVNEQAPTRGEAALLHYLDPD 480
DB 421 IYPGFGGAIIVFFMSDFPIAHGTNGLSVPCCTIPOEFVTHFVNEQAPTRGEAALLHYLDPD 480
QY 481 THRLNGEFLKYPEGFMTCVPNSSGTGPTLPINGVVFVSVSRFYQLKPVGTAGPACRL 540
DB 481 THRLNGEFLKYPEGFMTCVPNSSGTGPTLPINGVVFVSVSRFYQLKPVGTAGPACRL 540
QY 541 GIRRS 545
DB 541 GIRRS 545

RESULT 4
QY18C2 PRELIMINARY; PRT; 545 AA.
AC QY18C2;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Little Rock/316/1994/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Little Rock/316/1994/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glaes R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Little Rock/316/1994/US;
RX MEDLINE=96071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
```

```
Seto Y., Monroe S.S., Glaes R.I.;
"Correlation of patient immune responses with genetically
characterized small round-structured viruses involved in outbreaks of
nonbacterial acute gastroenteritis in the United States, 1990 to
1995.";
J. Med. Virol. 53:372-383(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Little Rock/316/1994/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Little Rock/316/1994/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glaes R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
EMBL; AF414405; AAL12968.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 545 AA; 5866 MW; F55076D61D612E82 CRC64;

Query Match      98.1%; Score 2840; DB 12; Length 545;
Best Local Similarity 98.5%; Pred. No. 1.4e-227;
Matches 537; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MMASKDPTNMDTSGAGQLVPEANTAEPISEMPVAGAAATAAGVNNIDPWIMNY 60
DB 1 MMASKDPTNMDTSGAGQLVPEANTAEPISEMPVAGAAATAAGVNNIDPWIMNY 60
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHNPFSLHQAQYNGWGNMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHNPFSLHQAQYNGWGNMKVKVLLAGNAFTA 120
QY 121 GKIIISCIIPGFAAQNISIAQATMPHVIADRVLEPIEVEPLEDVNRNVLFRNNNDNPTMR 180
DB 121 GKIIISCIIPGFAAQNISIAQATMPHVIADRVLEPIEVEPLEDVNRNVLFRNNNDNPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPPVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSVPNLPLN 240
DB 181 LVCMLYTPLRASGSSGTDPPVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSVPNLPLN 240
QY 241 TLSNRVPSLTKSMVSRDHQGMVQFNGRVTLTGQGLTPTTSASQLCKIRGSVFHANG 300
DB 241 TLSNRVPSLTKSMVSRDHQGMVQFNGRVTLTGQGLTPTTSASQLCKIRGSVFHANG 300
QY 301 GNGYNLTDLGSPYHAFESPAPIGFPDLGECDMHMEASPTTQFNTGDIKQINVKQESAF 360
DB 301 GNGYNLTDLGSPYHAFESPAPIGFPDLGECDMHMEASPTTQFNTGDIKQINVKQESAF 360
QY 361 APHLGTIOADGLSDVSVNTNMIKLGWVSPVSDGHRGVDVDPWIPRYGSTITEAAQLAPP 420
DB 361 APHLGTIOADGLSDVSVNTNMIKLGWVSPVSDGHRGVDVDPWIPRYGSTITEAAQLAPP 420
QY 421 IYPGFGGAIIVFFMSDFPIAHGTNGLSVPCCTIPOEFVTHFVNEQAPTRGEAALLHYLDPD 480
DB 421 IYPGFGGAIIVFFMSDFPIAHGTNGLSVPCCTIPOEFVTHFVNEQAPTRGEAALLHYLDPD 480
QY 481 THRLNGEFLKYPEGFMTCVPNSSGTGPTLPINGVVFVSVSRFYQLKPVGTAGPACRL 540
DB 481 THRLNGEFLKYPEGFMTCVPNSSGTGPTLPINGVVFVSVSRFYQLKPVGTAGPACRL 540
QY 541 GIRRS 545
DB 541 GIRRS 545

RESULT 5
QY1H12 PRELIMINARY; PRT; 545 AA.
ID QY1H12
AC QY1H12;
```

01-DEC-2001 (TrEMBLrel. 19, Created)
 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 Capsid.
 Human calicivirus NLV/VA98115/1998.
 Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 NCBI_TaxID=165509;
 [1]
 SEQUENCE FROM N.A.
 STRAIN=NLV/VA98115/1998;
 RC Jiang X., Zhong W.M., Farkas T., Huang P.W., Wilton N., Barrett E.,
 RA Fulton D., Morrow R., Matson D.O.;
 RT "Baculovirus expression and antigenic characterization of the capsid
 proteins of three Norwalk-like viruses.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY038598; AAK4673.1; -;
 DR InterPro; IPR004005; Calici_coat.
 DR Pfam; PF00915; Viral_Cap_coat.
 DR SEQUENCE 545 AA; 58295 MW; 258B0C7DEDB34DAC CRC64;
 SQ
 Query Match 96.6%; Score 2798; DB 12; Length 545;
 Best Local Similarity 96.7%; Pred. No. 4.2e-224; Indels 0; Gaps 0;
 Matches 527; Conservative 7; Mismatches 11;
 QY 1 MMASKADPTNMDGTSGAGQLVPEANTAEPISEMPVAGAAATAAATAGQVNMIDPWIMNYY 60
 DB 1 MMASKADPTNMDGTSGAGQLVPEANTAEPISEMPVAGAAATAAATAGQVNMIDPWIMNYY 60
 QY 61 VQAPQGEFTISPNPTGDIILFDLQGLPHLPFLSHLAQYNGWGMKVKVLLAGNAFTA 120
 DB 61 VQAPQGEFTISPNPTGDIILFDLQGLPHLPFLSHLAQYNGWGMKVKVLLAGNAFTA 120
 QY 121 GKIIISCIIPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVNRNLFHNNDNAPTMR 180
 DB 121 GKIIISCIIPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVNRNLFHNNDNAPTMR 180
 QY 181 LVCMLYTLRASGSSSGTDPFVIAGRLVTCPSDFSLFLVPPNVEQTKPFSVNLPLN 240
 DB 181 LVCMLYTLRASGSSSGTDPFVIAGRLVTCPSDFSLFLVPPNVEQTKPFSVNLPLN 240
 QY 241 TLNSRVPSLISKMSVSRDHGMQVQFNGRVTLDGQLQGTTPTSASOLCKIRGSVFHANG 300
 DB 241 TLNSRVPSLIRMSVSRDHGMQVQFNGRVTLDGQLQGTTPTSASOLCKIRGSVFHANG 300
 QY 301 GNGYLTDELGSPYHAFESPAPIGFDLGECDWHMEASPTTQFNTGDVTKQINVKQESAF 360
 DB 301 GNGYLTDELGSPYHAFESPAPIGFDLGECDWHMEASPTTQFNTGDVTKQINVKQESAF 360
 QY 361 APHLGTTQADGLSDVSNTNMIKLGWSPVSDGHRGVDVDPWIPRYGSTLTEAAQLAPP 420
 DB 361 APHLGTTQADGLSDVSNTNMIKLGWSPVSDGHRGVDVDPWIPRYGSTLTEAAQLAPP 420
 QY 421 IYPGFGAEIVFMSDFPIAHGNTGLSVPTCTIQEFVTHFVNEQAPTRGEAALLHYLDDP 480
 DB 421 IYPGFGAEIVFMSDFPIAHGNTGLSVPTCTIQEFVTHFVNEQAPTRGEAALLHYLDDP 480
 QY 481 THRNLFGEFLKLYPEGFMTCVPSNSGTGQTLPIGNGVFWVSWSRFYQLKPVGTAGPACRL 540
 DB 481 THRNLFGEFLKLYPEGFMTCVPSNSGTGQTLPIGNGVFWVSWSRFYQLKPVGTAGPACRL 540
 QY 541 GIRRS 545
 DB 541 GIRRS 545
 RESULT 6
 Q91V45 PRELIMINARY; PRT; 545 AA.
 ID Q91V45
 AC Q91V45;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 Capsid protein.
 Human calicivirus HU/NLV/Birmingham/93/UK.
 Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 NCBI_TaxID=122916;
 [1]
 SEQUENCE FROM N.A.
 STRAIN=HU/NLV/Birmingham/93/UK;
 RC MEDLINE=20404883; PubMed=10949950;
 RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
 RA Clegg J.C., Chamberlain J., Brown D.W.G.;
 RT "Capsid protein diversity among 'Norwalk-like' viruses.";
 RL Virus Genes 20:227-236(2000).
 DR EMBL; AU277612; CAB89093.1; -;
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_Cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 DR SEQUENCE 545 AA; 58689 MW; DC6EBC6EBEBCDF7C7 CRC64;
 SQ
 Query Match 94.8%; Score 2744; DB 12; Length 545;
 Best Local Similarity 95.8%; Pred. No. 1.3e-219; Indels 0; Gaps 0;
 Matches 522; Conservative 7; Mismatches 16;
 QY 1 MMASKADPTNMDGTSGAGQLVPEANTAEPISEMPVAGAAATAAATAGQVNMIDPWIMNYY 60
 DB 1 MMASKADPTNMDGTSGAGQLVPEANTAEPISEMPVAGAAATAAATAGQVNMIDPWIMNYY 60
 QY 61 VQAPQGEFTISPNPTGDIILFDLQGLPHLPFLSHLAQYNGWGMKVKVLLAGNAFTA 120
 DB 61 VQAPQGEFTISPNPTGDIILFDLQGLPHLPFLSHLAQYNGWGMKVKVLLAGNAFTA 120
 QY 121 GKIIISCIIPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVNRNLFHNNDNAPTMR 180
 DB 121 GKIIISCIIPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVNRNLFHNNDNAPTMR 180
 QY 181 LVCMLYTLRASGSSSGTDPFVIAGRLVTCPSDFSLFLVPPNVEQTKPFSVNLPLN 240
 DB 181 LVCMLYTLRASGSSSGTDPFVIAGRLVTCPSDFSLFLVPPNVEQTKPFSVNLPLN 240
 QY 241 TLNSRVPSLISKMSVSRDHGMQVQFNGRVTLDGQLQGTTPTSASOLCKIRGSVFHANG 300
 DB 241 TLNSRVPSLISKMSVSRDHGMQVQFNGRVTLDGQLQGTTPTSASOLCKIRGSVFHANG 300
 QY 301 GNGYLTDELGSPYHAFESPAPIGFDLGECDWHMEASPTTQFNTGDVTKQINVKQESAF 360
 DB 301 GNGYLTDELGSPYHAFESPAPIGFDLGECDWHMEASPTTQFNTGDVTKQINVKQESAF 360
 QY 361 APHLGTTQADGLSDVSNTNMIKLGWSPVSDGHRGVDVDPWIPRYGSTLTEAAQLAPP 420
 DB 361 APHLGTTQADGLSDVSNTNMIKLGWSPVSDGHRGVDVDPWIPRYGSTLTEAAQLAPP 420
 QY 421 IYPGFGAEIVFMSDFPIAHGNTGLSVPTCTIQEFVTHFVNEQAPTRGEAALLHYLDDP 480
 DB 421 IYPGFGAEIVFMSDFPIAHGNTGLSVPTCTIQEFVTHFVNEQAPTRGEAALLHYLDDP 480
 QY 481 THRNLFGEFLKLYPEGFMTCVPSNSGTGQTLPIGNGVFWVSWSRFYQLKPVGTAGPACRL 540
 DB 481 THRNLFGEFLKLYPEGFMTCVPSNSGTGQTLPIGNGVFWVSWSRFYQLKPVGTAGPACRL 540
 QY 541 GIRRS 545
 DB 541 GIRRS 545
 RESULT 7
 Q66418 PRELIMINARY; PRT; 544 AA.
 ID Q66418
 AC Q66418;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Capsid protein.

OS Desert Shield virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 OX NCBI_TaxID=33755;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSV395;
 RX MEDLINE=94174735; PubMed=8128635;
 RA Lew J.F., Kapikian A.Z., Jiang X., Estes M.K., Green K.Y.;
 RT "Molecular characterization and expression of the capsid protein of a
 Norwalk-like virus recovered from a Desert Shield troop with
 gastroenteritis.";
 RL J. Virol. 77:2004-2009, 2003.
 DR EMBL; U04469; AAA16285.1; -
 DR InterPro; IPR004005; Calici.coat.
 DR Pfam; PF00915; Calici.coat.1.
 SQ SEQUENCE 544 AA; 58606 MW; 07A4AE74BC581E8 CRC64;

Query Match 90.0%; Score 2605.5; DB 12; Length 544;
 Best Local Similarity 89.2%; Pred. No. 4.4e-208;
 Matches 486; Conservative 24; Mismatches 34; Indels 1; Gaps 1;

Qy 1 MMASKDAPTNDGTSGAGQLVPEANTAEPI SMEPVAGAAATAAGVNNIDPWIMNY 60
 Db 1 MMASKDAPTNDGTSGAGQLVPEANTAEPI SMEPVAGAAATAAGVNNIDPWIMNSY 60

Qy 61 VOAPQGETTISPNTPGDILFDLQGLPHNLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120
 Db 61 VOAPQGETTISPNTPGDILFDLQGLPHNLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120

Qy 121 GKIIISCIPTPGFAAQNISIAQATWPHVIAVDRVLEPIEVLDPVNNVLFHNDNAPTMR 180
 Db 121 GKIIISCIPTPGFAAQNISIAQATWPHVIAVDRVLEPIEVLDPVNNVLFHNDNAPTMR 180

Qy 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCTSPDFSLFLVPPNVQKTKPFSVPLNPLN 240
 Db 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCTSPDFSLFLVPPNVQKTKPFSVPLNPLN 240

Qy 241 TLSNRVPSLTKSMVSRDHQOMQVQFNGRVTLQGLQGTTPTSASQCKIRGVSFHANG 300
 Db 241 TLSNRVPSLTKSMVSRDHQOMQVQFNGRVTLQGLQGTTPTSASQCKIRGVSFHANG 300

Qy 301 GNGYNLTLDGSPYHAFESPAIPGPDLDGECDDHMEASPTTFQNTGDIKQINVKQESAF 360
 Db 301 GNGYNLTLDGSPYHAFESPAIPGPDLDGECDDHMEASPTTFQNTGDIKQINVKQESAF 360

Qy 361 APHLGTVQADNLS-AGANTDLIVSLWSISVSDQHRHDVDPWVIPRYGSSLTEAAQLAPP 420
 Db 361 APHLGTVQADNLS-AGANTDLIVSLWSISVSDQHRHDVDPWVIPRYGSSLTEAAQLAPP 420

Qy 421 IYPGFGGAIYFFMSDFPIAGTNGLSVPCTIPQEFVTHFVNEQAPTRGEAALLHYLDPD 480
 Db 421 IYPGFGGAIYFFMSDFPIAGTNGLSVPCTIPQEFVTHFVNEQAPTRGEAALLHYLDPD 480

Qy 481 THRLNGEFLKYPEGFMTCVPNSSGTGPTLPINGVVFVSVMSVSRFYQLKPVGTAGPACRL 540
 Db 481 THRLNGEFLKYPEGFMTCVPNSSGTGPTLPINGVVFVSVMSVSRFYQLKPVGTAGPACRL 540

Qy 541 GIRRS 545
 Db 540 GIRRS 544

RESULT 8
 Q8BCA3 ID Q8BCA3 PRELIMINARY; PRT; 543 AA.
 AC Q8BCA3;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DE Capsid.
 OS Human calicivirus NLV/Boxer/2001/US.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 OX NCBI_TaxID=207658;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NLV/Boxer/2001/US;
 RA Farkas T., Thornton S.A., Zhong W., Jiang X.;
 RT "Viral gastroenteritis outbreaks on US Navy Ships.";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF538679; AAN15140.1; -
 DR InterPro; IPR004005; Calici.coat.
 DR InterPro; IPR008975; Viral_Cap_coat.
 DR Pfam; PF00915; Calici.coat.1.
 SQ SEQUENCE 543 AA; 58783 MW; 9CFD55A1858E0F62 CRC64;

Query Match 73.9%; Score 2139; DB 12; Length 543;
 Best Local Similarity 72.1%; Pred. No. 3e-169;
 Matches 397; Conservative 74; Mismatches 66; Indels 14; Gaps 7;

Qy 1 MMASKDAPTNDGTSGAGQLVPEANTAEPI SMEPVAGAAATAAGVNNIDPWIMNY 60
 Db 1 MMASKDAPTNDGTSGAGQLVPGANTAEPPIKPVAGAAATAAGVNNIDPWIMNNF 60

Qy 61 VOAPQGETTISPNTPGDILFDLQGLPHNLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120
 Db 61 VOAPQGETTISPNTPGDILFDLQGLPHNLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120

Qy 121 GKIIISCIPTPGFAAQNISIAQATWPHVIAVDRVLEPIEVLDPVNNVLFHNDNAPTMR 180
 Db 121 GKIIISCIPTPGFAAQNISIAQATWPHVIAVDRVLEPIEVLDPVNNVLFHNDNAPTMR 180

Qy 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCTSPDFSLFLVPPNVQKTKPFSVPLNPLN 240
 Db 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCTSPDFSLFLVPPNVQKTKPFSVPLNPLN 240

Qy 241 TLSNRVPSLTKSMVSRDHQOMQVQFNGRVTLQGLQGTTPTSASQCKIRGVSFHANG 300
 Db 241 TLSNRVPSLTKSMVSRDHQOMQVQFNGRVTLQGLQGTTPTSASQCKIRGVSFHANG 300

Qy 301 GNGYNLTLDGSPYHAFESPAIPGPDLDGECDDHMEASPTTFQNTGDIKQINVKQESAF 356
 Db 301 GNGYNLTLDGSPYHAFESPAIPGPDLDGECDDHMEASPTTFQNTGDIKQINVKQESAF 356

Qy 357 ESAPFAPLGTITQADGLSDVS--VNTNMIKLGWSPVSDGHRGVDVDPWVIPRYGSSLTEA 414
 Db 357 ESAPFAPLGTITQADGLSDVS--VNTNMIKLGWSPVSDGHRGVDVDPWVIPRYGSSLTEA 414

Qy 360 DATFAPHLGTVKLEDDNNELDQFVGEVLELTW---VSNRTGATLNLWAVPNYGSNLQA 416
 Db 360 DATFAPHLGTVKLEDDNNELDQFVGEVLELTW---VSNRTGATLNLWAVPNYGSNLQA 416

Qy 415 AQLAPPYPPGFGGAIYFFMSDFPIAGTNGLSVPCTIPQEFVTHFVNEQAPTRGEAALL 474
 Db 415 AQLAPPYPPGFGGAIYFFMSDFPIAGTNGLSVPCTIPQEFVTHFVNEQAPTRGEAALL 474

Qy 417 SQLAPPYPPGFGGAIYFFMSDFPIAGTNGLSVPCTIPQEFVTHFVNEQAPTRGEAALL 473
 Db 417 SQLAPPYPPGFGGAIYFFMSDFPIAGTNGLSVPCTIPQEFVTHFVNEQAPTRGEAALL 473

Qy 475 HYLPDPTHRLNGEFLKYPEGFMTCVPNSSGTGPTLPINGVVFVSVMSVSRFYQLKPVGTA 534
 Db 475 HYLPDPTHRLNGEFLKYPEGFMTCVPNSSGTGPTLPINGVVFVSVMSVSRFYQLKPVGTA 534

Qy 535 GPACRLGIRRS 545
 Db 533 GAARLGLRRS 543

RESULT 9
 Q9IV48 ID Q9IV48 PRELIMINARY; PRT; 539 AA.
 AC Q9IV48;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE Capsid protein.
 OS Human calicivirus HU/NLV/Winchester/94/UK.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 OX NCBI_TaxID=122913;
 RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/Winchester/94/UK;
RX MEDLINE=20404883; PubMed=10949950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL; AJ277609; CAB89090.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 539 AA; 58120 MW; 14E612D7A576395D CRC64;

Query Match 73.3%; Score 2124; DB 12; Length 539;
Best Local Similarity 72.7%; Pred. No. 5, 2e-168;
Matches 397; Conservative 56; Mismatches 83; Indels 10; Gaps 6;

QY 1 MMMSKDAFTNNMDGTSAGQLVPEANTAEPISEMPEVAGAAATAAGQVNMIDPWIMNNY 60
DB 1 MMMSKDVPSNMDXSGAGQLVPEVNAEPLPLEPVGAATAAATAAGQVNLIDPWIMNF 60

QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120

QY 121 GKIISCIPIPPGFAAQNISIAQATMPHVIADVRVLEPIEVPLEDVRNVLFNNDVPTMR 180
DB 121 GKIICCVPPGASQNISIGQATMPHVIADVRVLEPIEIPDDVRNVLFTNENRPTMR 180

QY 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFVSVPNLPLN 240
DB 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFVSVPNLPLN 240

QY 241 TLNSRVPSLIKMSVSRDHQGMVQFNGRVTLTGQGLTTPTSASQLCKIRGSVFHANG 300
DB 241 NLANSRVPMINKVTVSTQNVQVQFNGRCILEGQLLTTPVSASQVARIKGVFSTAS 300

QY 301 GNGYNLTDLGSPYHAFESPAPTFGPDLCEDWHM--EASPTQNTGTVIKQINVKOES 358
DB 301 GKGLNLTDLGTPYHAFESPAPTFGPDLCEDWHM--EASPTQNTGTVIKQINVKOES 358

QY 359 AFPHLGTIQADGLSDVSVNNTMIKLGWVSPVSDGHRGVDVDPWVIPRYGSLTTEAAQLA 418
DB 359 PICTXGLSIEF--TSDDQDAGDQLGTLAWVSPSTSGAR--VDPWKIPSYGSTVTESTILA 414

QY 419 PPIYPFGGEAIVFMSDFPIAHGTNGLSVPTCTIQEFVTHFVNEQAPRGEAALLHYLD 478
DB 415 PLIFPPGGEAIVFMSDFPIVSG-NTAQIPCTLPQEFVSHFVEQQAPINGEAALLHYVD 473

QY 479 PDTHRNLFGEFLYPEGFWMTCPVNSSGTPQTLPIGVFVSVWSRYQKPKVGTAGPAC 538
DB 474 PDTHRNLFGEFLYPDGFITCVNTGG-GPQNLFTGVFVSVWSRYQKPKVGTAGPAR 532

QY 539 RLGIIR 544
DB 533 RLGVRR 538

RESULT 10
Q8XJJO PRELIMINARY; PRT; 543 AA.
AC Q8XJJO;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid.
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SzUG1;

```

```

RA Katayama K., Kojima S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RC STRAIN=SzUG1;
RX MEDLINE=22192455; PubMed=12202225;
RA Katayama K., Shirato-Horikoshi H., Kojima S., Kageyama T., Oka T.,
RA Hoshino F.B., Fukushi S., Shinohara M., Uchida K., Suzuki Y.,
RA Gojobori T., Takeda N.;
RT "Phylogenetic Analysis of the Complete Genome of 18 Norwalk-like
RT Viruses.";
RL Virology 299:225-239(2002).
DR EMBL; AB039774; BAC11813.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 543 AA; 58761 MW; E0DE7489A9C3488C CRC64;

Query Match 69.7%; Score 2017.5; DB 12; Length 543;
Best Local Similarity 70.1%; Pred. No. 3, 9e-159;
Matches 385; Conservative 58; Mismatches 95; Indels 11; Gaps 7;

QY 1 MMMSKDAFTNNMDGTSAGQLVPEANTAEPISEMPEVAGAAATAAGQVNMIDPWIMNNY 60
DB 1 MMMSKDATPSADGANGAGQLVPEVNAEPLPLDPVAGASATATAGQVNMIDPWIMNF 60

QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120

QY 121 GKIISCIPIPPGFAAQNISIAQATMPHVIADVRVLEPIEVPLEDVRNVLFNNDVPTMR 180
DB 121 GKIICCVPPGFSRSLTSAQATLFPHVIVADVRVLEPIEIPLEDVRNVLHNNDVPTMR 180

QY 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFVSVPNLPLN 240
DB 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFVSVPNLPLN 240

QY 241 TLNSRVPSLIKMSVSRDHQGMVQFNGRVTLTGQGLTTPTSASQLCKIRGSVFHANG 300
DB 241 TLNSRVPSLIKMSVSRDHQGMVQFNGRVTLTGQGLTTPTSASQLCKIRGSVFHANG 300

QY 301 GNGYNLTDLGSPYHAFESPAPTFGPDLCEDWHM--EASPTQNTGTVIKQINVKOES 359
DB 301 GNGYNLTDLGSPYHAFESPAPTFGPDLCEDWHM--EASPTQNTGTVIKQINVKOES 359

QY 359 AFPHLGTIQADGLSDVSVNNTMIKLGWVSPVSDGHRGVDVDPWVIPRYGSLTTEAAQL 417
DB 360 FAPHLGTIQADGLSDVSVNNTMIKLGWVSPVSDGHRGVDVDPWVIPRYGSLTTEAAQL 416

QY 419 PPIYPFGGEAIVFMSDFPIAHGTNGLSVPTCTIQEFVTHFVNEQAPRGEAALLHYLD 478
DB 415 PLIFPPGGEAIVFMSDFPIVSG-NTAQIPCTLPQEFVSHFVEQQAPINGEAALLHYVD 476

QY 479 PDTHRNLFGEFLYPEGFWMTCPVNSSGTPQTLPIGVFVSVWSRYQKPKVGTAGPAC 538
DB 474 PDTHRNLFGEFLYPDGFITCVNTGG-GPQNLFTGVFVSVWSRYQKPKVGTAGPAR 532

QY 539 RLGIIR 544
DB 533 RLGVRR 538

RESULT 11
Q9DU46 PRELIMINARY; PRT; 544 AA.
AC Q9DU46;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Chiba virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;

```

```
OC Norovirus.
OX NCBI_TaxID=99565;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Hu/NL/Chiba 407/1987/Jp;
RA Soneya Y., Takeda N., Miyamura T.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=Hu/NL/Chiba 407/1987/Jp;
RX MEDLINE=20569531; PubMed=11118371;
RA Soneya Y., Takeda N., Miyamura T.;
RT "Complete nucleotide sequence of the Chiba virus genome and functional
RT expression of the 3C-like protease in Escherichia coli.";
RL Virology 278:490-500(2000).
DR EMBL; AB042808; BAB18267.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58351 MW; 1DB7A4E81FBB246F CRC64;

Query Match 68.3%; Score 1979; DB 12; Length 544;
Best Local Similarity 66.8%; Pred. No. 6.2e-156;
Matches 368; Conservative 68; Mismatches 101; Indels 14; Gaps 6;

Qy 1 MMASKADPTNMDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGVNNIDPWIMNYY 60
Db 1 MMASKADPTSGADGAGQLVPEVNTADPIPDPVAGSSTALATAGVNLDPWIINNF 60
Qy 61 VQAPQGEFTISPNNTPGDILFDLQGLHPLNPFSLHQAQMYNGWGNKVKVLLAGNAFTA 120
Db 61 VQAPQGEFTISPNNTPGDVLFDLQGLHPLNPFSLHQAQMYNGWGNKVRVVLAGNAFTA 120
Qy 121 GKIIISICPPGFAAQNISIAQATMFPHVIADRVLEPIEVLPELDVNRVLFHNNDNAPTMR 180
Db 121 GKVIICVPPGFQRTLSIAQATLFPHVIAADVRLDPEVLELDVNRVLYHNNDTQPTMR 180
Qy 181 LVCMLYTLPLRASGSGGTPDFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPVNLPLN 240
Db 181 LLCMLYTLPLRTGGASGGTDSFVAGRVLTCPGDFNLFVPPVTEQKTRPFTVPIPLK 240
Qy 241 TLSNRVPSLIKSMVSRDHQMVQFQNGRVLTLDGQLQGTTPTSASQLCKIRGSVFHANG 300
Db 241 YLSNRIPNPIEGMSLSPDQTNQVQFQNGRCLTDGQPLGTTTPVSVSQLCKFRGRI--TSG 298
Qy 301 GNGVNLTELDGSPVHAFESPAIGFDPDLGECDHMEAS--PTTQFNTGDVVKQINVKQES 358
Db 299 QRVNLTELDGSPFMAFAAPAGFPDLGSCDHMEISKIPNSQTQNPVITNSVKPNSQ 358
Qy 359 AFAPHLGTIQADGLSDVSNTNMIKLGWSPVSDGHRGDVDPWVI PRYGSTLTEAAQLA 418
Db 359 QFVPHLSSITLD--ENVSSGGDIYGTQWTSPPSDSGGANTNFKIPDYGSSLAESAQLA 416
Qy 419 PPIYPGFGGAIVFFMSDFPIAHGTNGLS-----VPCTIPQEFVTHFVNEQAPTGEAALL 474
Db 417 PAVYPPGFNEVIVFMAISIP---GPNQSGSNLVPCLLPQBYITHFISEQAPIQGEAALL 473
Qy 475 HYLDPDTHRLNGLGEFKLYPEGFMTCVPNSSGTPQTLPINGVVFVSVWSRYQLKPVGTA 534
Db 474 HYVDPDTHRLNGLGEFKLYPGGYLTCPVNSSSTGTPQPLDGLGVFVFSWVSRFYQLKPVGTA 533
Qy 535 GPA-CRLGIRR 544
Db 534 GPARGRLGVR 544

RESULT 12
Q8JW44
ID Q8JW44 PRELIMINARY; PRT; 546 AA.
AC Q8JW44;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```

```
DE Capsid protein.
GN CAPSID.
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Funabashi258;
RA Katayama K., Takeda N., Natori K.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=Funabashi258;
RA Natori K., Takeda N.;
RT "Genetic and antigenic relationship among Norwalk-like viruses.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB078335; BAC05516.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 546 AA; 58834 MW; 1965F054E2C48186 CRC64;

Query Match 68.3%; Score 1977.5; DB 12; Length 546;
Best Local Similarity 66.6%; Pred. No. 8.3e-156;
Matches 367; Conservative 76; Mismatches 95; Indels 13; Gaps 7;

Qy 1 MMASKADPTNMDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGVNNIDPWIMNYY 60
Db 1 MMASKADPTSGADGAGQLVPEVNTADPIPDPVAGSSTALATAGVNLDPWIINNF 60
Qy 61 VQAPQGEFTISPNNTPGDILFDLQGLHPLNPFSLHQAQMYNGWGNKVKVLLAGNAFTA 120
Db 61 VQAPQGEFTISPNNTPGDILFDLQGLHPLNPFSLHQAQMYNGWGNKVRVVLAGNAFTA 120
Qy 121 GKIIISICPPGFAAQNISIAQATMFPHVIADRVLEPIEVLPELDVNRVLFHNNDNAPTMR 180
Db 121 GKVIICVPPGFQRTLSIAQATLFPHVIAADVRLDPEVLELDVNRVLYHNNDTQPTMR 180
Qy 181 LVCMLYTLPLRASGSGGTPDFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPVNLPLN 240
Db 181 LVCMLYTLPLRTGGASGGTDSFVAGRVLTCPGDFNLFVPPVTEQKTRPFTVPIPLQ 240
Qy 241 TLSNRVPSLIKSMVSRDHQMVQFQNGRVLTLDGQLQGTTPTSASQLCKIRGSVFHANG 300
Db 241 TLSNRFPFSLIQMILSPDASQVVFQNGRCLTDGQLGTTTPATSGQLFRVRGKI--NOG 298
Qy 301 GNGVNLTELDGSPVHAFESPAIGFDPDLGECDHMEASPT-TQFNTGDVVKQINVKQES-S 358
Db 299 ARTLNLTEDGSPFMAFAAPAGFPDLGSCDHMEISKIPNSQTQNPVITNSVKPNSQ 358
Qy 359 AFAPHLGTIQADGLSDVSNTNMIKLGWSPVSDGHRGDVDPWVI PRYGSTLTEAAQLA 418
Db 359 GFVPHLGSIQDFEVNHTPG-DYIGTIEWISQSPSTPGCTDNLNWEIPDYGSSLSQAANLA 417
Qy 419 PPIYPGFGGAIVFFMSDFPIAHGTNGLS-----VPCTIPQEFVTHFVNEQAPTGEAALL 474
Db 418 PPVPPGFGGALVYVVSAFP---GPNRSAPNDVPCLLPQBYITHFVSEQAPTNGDAALL 474
Qy 475 HYLDPDTHRLNGLGEFKLYPEGFMTCVPNSSGTPQTLPINGVVFVSVWSRYQLKPVGTA 534
Db 475 HYVDPDTHRLNGLGEFKLYPGGYLTCPVNGVGAGPQPLNGVFLFVSVWSRYQLKPVGTA 534
Qy 535 GPA-CRLGIRR 544
Db 535 STARGRLGVR 545

RESULT 13
Q91185
ID Q91185 PRELIMINARY; PRT; 544 AA.
AC Q91185;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
```

DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)	OS	Chiba virus.
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
DE	Capsid protein.	OC	Norovirus.
OS	Human calicivirus HU/NLV/Koblentz/433/2000/DE.	OX	NCBI_TaxID=99565;
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;	RN	[1]
OC	Norovirus.	RP	SEQUENCE FROM N.A.
OX	NCBI_TaxID=165252;	RC	STRAIN=Chiba 407;
RN	[1]	RA	Natori K., Suzuki K., Yamakawa Y., Tatsumi M., Sakae K., Kobayashi S.,
RP	SEQUENCE FROM N.A.	RA	Shinozaki K., Ishiko H., Miyamura T., Takeda N.;
RC	STRAIN=HU/NLV/Koblentz 433/2000/DE;	RT	"Expression and self-assembly of capsid proteins of the Chiba virus, a
RA	Kuenkel U., Schreier E.;	RT	genetically distinct Norwalk-like virus";
RT	"Molecular epidemiology of outbreaks of gastroenteritis associated	RL	Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RT	with Norwalk-like viruses in Germany.";	DR	EMBL; AB022679; BAA82106.1; -
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.	DR	InterPro; IPR004005; Calici.coat.
DR	EMBL; AF394960; AAK72048.1; -	DR	InterPro; IPR008975; Viral_cap_coat.
DR	InterPro; IPR004005; Calici.coat.	DR	Pfam; PF00915; Calici_coat_1.
DR	InterPro; IPR008975; Viral_cap_coat.	SQ	SEQUENCE 544 AA; 58323 MW; ABDLC1FC4F93D872 CRC64;
DR	Pfam; PF00915; Calici_coat_1.		
SQ	SEQUENCE 544 AA; 58222 MW; 1B39AS95B733A6A8 CRC64;		
Query Match 68.2%; Score 1975; DB 12; Length 544;			
Best Local Similarity 66.8%; Pred. No. 1.3e-155;			
Matches 368; Conservative 68; Mismatches 101; Indels 14; Gaps 6;			
Qy	1 MMASKDAPTNDGTSAGOLVPEANTAEPISEMPVAGATAAATAGOVNMDPWIMNYY 60	Qy	1 MMASKDAPTNDGTSAGOLVPEANTAEPISEMPVAGATAAATAGOVNMDPWIMNYY 60
Db	1 MMASKDAPTNDGTSAGOLVPEANTAEPISEMPVAGATAAATAGOVNMDPWIMNYY 60	Db	1 MMASKDAPTNDGTSAGOLVPEANTAEPISEMPVAGATAAATAGOVNMDPWIMNYY 60
Qy	61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120	Qy	61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
Db	61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120	Db	61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
Qy	121 GKIIISCIIPGFAAQNISIAQATMFPHVIAADVRLPEVPLEDVRNVLPHNNDNAPTMR 180	Qy	121 GKIIISCIIPGFAAQNISIAQATMFPHVIAADVRLPEVPLEDVRNVLPHNNDNAPTMR 180
Db	121 GKIIISCIIPGFAAQNISIAQATMFPHVIAADVRLPEVPLEDVRNVLPHNNDNAPTMR 180	Db	121 GKIIISCIIPGFAAQNISIAQATMFPHVIAADVRLPEVPLEDVRNVLPHNNDNAPTMR 180
Qy	181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFVPPNVEQTKPFSPVNLPLN 240	Qy	181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFVPPNVEQTKPFSPVNLPLN 240
Db	181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFVPPNVEQTKPFSPVNLPLN 240	Db	181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFVPPNVEQTKPFSPVNLPLN 240
Qy	241 TLNSRVPSLIKSMVSRDHGQVQFQNGRVTLTGQGLQGTTPTSASOLCKIRGSVFHANG 300	Qy	241 TLNSRVPSLIKSMVSRDHGQVQFQNGRVTLTGQGLQGTTPTSASOLCKIRGSVFHANG 300
Db	241 TLNSRVPSLIKSMVSRDHGQVQFQNGRVTLTGQGLQGTTPTSASOLCKIRGSVFHANG 300	Db	241 TLNSRVPSLIKSMVSRDHGQVQFQNGRVTLTGQGLQGTTPTSASOLCKIRGSVFHANG 300
Qy	301 GNGYNLTDLGSPYHAFESAPATGFDLGECDHMEAS--PTQFNTGDIKQINVKQES 358	Qy	301 GNGYNLTDLGSPYHAFESAPATGFDLGECDHMEAS--PTQFNTGDIKQINVKQES 358
Db	301 GNGYNLTDLGSPYHAFESAPATGFDLGECDHMEAS--PTQFNTGDIKQINVKQES 358	Db	301 GNGYNLTDLGSPYHAFESAPATGFDLGECDHMEAS--PTQFNTGDIKQINVKQES 358
Qy	358 QKVLNLTDLGSPYHAFESAPATGFDLGECDHMEAS--PTQFNTGDIKQINVKQES 358	Qy	358 QKVLNLTDLGSPYHAFESAPATGFDLGECDHMEAS--PTQFNTGDIKQINVKQES 358
Db	358 QKVLNLTDLGSPYHAFESAPATGFDLGECDHMEAS--PTQFNTGDIKQINVKQES 358	Db	358 QKVLNLTDLGSPYHAFESAPATGFDLGECDHMEAS--PTQFNTGDIKQINVKQES 358
Qy	359 AFAPHLGTIQADGLSDVSVNNTMIAGLVSPVSDGHRGVDVDPWVPIRYGSLTTEAAQLA 418	Qy	359 AFAPHLGTIQADGLSDVSVNNTMIAGLVSPVSDGHRGVDVDPWVPIRYGSLTTEAAQLA 418
Db	359 AFAPHLGTIQADGLSDVSVNNTMIAGLVSPVSDGHRGVDVDPWVPIRYGSLTTEAAQLA 418	Db	359 AFAPHLGTIQADGLSDVSVNNTMIAGLVSPVSDGHRGVDVDPWVPIRYGSLTTEAAQLA 418
Qy	418 QFVPHLSSITLD--DNVSSGGDYIGTQWTSPPSDSGGANTNPKIPDYGSSLAESAQLA 416	Qy	418 QFVPHLSSITLD--DNVSSGGDYIGTQWTSPPSDSGGANTNPKIPDYGSSLAESAQLA 416
Db	418 QFVPHLSSITLD--DNVSSGGDYIGTQWTSPPSDSGGANTNPKIPDYGSSLAESAQLA 416	Db	418 QFVPHLSSITLD--DNVSSGGDYIGTQWTSPPSDSGGANTNPKIPDYGSSLAESAQLA 416
Qy	419 PPIYPPGGEALVFFMSDFPIAHGTNGLS----VPCTIPEQFVTHFVNEQAPRGEAALL 474	Qy	419 PPIYPPGGEALVFFMSDFPIAHGTNGLS----VPCTIPEQFVTHFVNEQAPRGEAALL 474
Db	419 PPIYPPGGEALVFFMSDFPIAHGTNGLS----VPCTIPEQFVTHFVNEQAPRGEAALL 474	Db	419 PPIYPPGGEALVFFMSDFPIAHGTNGLS----VPCTIPEQFVTHFVNEQAPRGEAALL 474
Qy	473 PAVYPPGGEALVFFMSDFPIAHGTNGLS----VPCTIPEQFVTHFVNEQAPRGEAALL 473	Qy	473 PAVYPPGGEALVFFMSDFPIAHGTNGLS----VPCTIPEQFVTHFVNEQAPRGEAALL 473
Db	473 PAVYPPGGEALVFFMSDFPIAHGTNGLS----VPCTIPEQFVTHFVNEQAPRGEAALL 473	Db	473 PAVYPPGGEALVFFMSDFPIAHGTNGLS----VPCTIPEQFVTHFVNEQAPRGEAALL 473
Qy	475 HYLPDPTNRNLGEFLKLYPEGMTCVPSNSTGTPQLPFGVVFVSWRSFYQLKPVGTA 534	Qy	475 HYLPDPTNRNLGEFLKLYPEGMTCVPSNSTGTPQLPFGVVFVSWRSFYQLKPVGTA 534
Db	475 HYLPDPTNRNLGEFLKLYPEGMTCVPSNSTGTPQLPFGVVFVSWRSFYQLKPVGTA 534	Db	475 HYLPDPTNRNLGEFLKLYPEGMTCVPSNSTGTPQLPFGVVFVSWRSFYQLKPVGTA 534
Qy	534 GPA-CRLGIRR 544	Qy	534 GPA-CRLGIRR 544
Db	534 GPARGLGVRR 544	Db	534 GPARGLGVRR 544
RESULT 14			
Q9QTE7 PRELIMINARY; PRT; 544 AA.			
ID	Q9QTE7	ID	Q9QTE7
AC	Q9QTE7	AC	Q9QTE7
DT	01-MAY-2000 (TrEMBLrel. 13, Created)	DT	01-MAY-2000 (TrEMBLrel. 13, Created)
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)	DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Capsid protein.	DE	Capsid protein.

```
OC Norovirus.
OX NCBI_TaxID=11983;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Norwalk-like;
RX MEDLINE=20254531; PubMed=10795514;
RA Schreier E., Doering F., Kuenkel U.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated
RL with small round structured viruses in Germany in 1997/98.";
RN Arch. Virol. 145:443-453(2000).
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Norwalk-like;
RA Schreier E., Doering F., Kuenkel U.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF093797; AAC64603.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 540 AA; 58018 MW; 647843487A654CBE CRC64;

Query Match      68.1%; Score 1972.5; DB 12; Length 540;
Best Local Similarity 67.7%; Pred. No. 2.1e-155;
Matches 377; Conservative 60; Mismatches 89; Indels 31; Gaps 7;

Qy 1 MMASKDAPTNDGTSGAGQLVPEANTABPISMEPVAGAATAAATAGQVNMIDPWIMNY 60
Db 1 MMASKDAPSPDGASGAGQLVPEVTADQISMDPVAGASTAVATAGQVNMIDPWIFNF 60

Qy 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQYNGVGNMKVKVLLAGNAFTA 120
Db 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLAHLNQMYNGVGNMRVRIILAGNAFTA 120

Qy 121 GKIIISCTIPCGFAAQTISIAQATMFPHVIAADVRLPIEVLPLEDRVNLPHNNDN-APT 179
Db 121 GKVIICVPPGFDARILHIAQTILFPHLIADVRTLEPVELEDRVNLHNSQPOPTM 180

Qy 180 RLVCMLYTLPRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQTKXPSVPLPL 239
Db 181 RLVALMYTLPLTGGSGGTDAFVAGRVLTCPAPDFSLFLVPPSVEQKTRVFSVPL 240

Qy 240 NTLNSRVPSLIKSMVSRDHQVQFQNGRVTLGDLQGTTPTSASQLCKIRGSVFHAN 299
Db 241 KDLNSRVPTLIQGMFVSPDYNQSVQFQNGRCQIDGQLGTTVPVLSQLCKIRGT--SS 298

Qy 300 GGNGYNLTLDGSPVHAFESAPIGFPLDGECDWHM-----EASPTTQENTGVK 350
Db 299 NTRVLNLSEVDGTPFVPLESPAPVGFDPDGGCDWHVGTTFEARDQDPSONVTFATND--- 355

Qy 351 QINVKQESAFAPHLGTI---QADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPWVIPRY 407
Db 356 -----SSFVPYLGSIAPHNGDGFH---SGDIIIGSLDWISAPSDGSALDV--WSIPKY 402

Qy 408 GSTLTERAAQLAPPYIPGFGGAIIVFFMSDFPIANGTNGLSVPCTIPQEFVTHFVNEOAPT 467
Db 403 GSSLPDVTHLAPAVFPFGFGEVILYFHSKFGSGPTDKLRVPCLIPQEFITHFCNEQAPI 462

Qy 468 RGEAALHYLPDTHRNIGEPKLYPEGFMTCVPNSSGTGQOTLPINGVFVSVWSRFYQ 527
Db 463 AGEAALHYVDPDTHRNIGEPKLYPDGFMTCVPNSISGGPQTLPIGVFVSVWSRFYQ 522

Qy 528 LKPVGTAGPACRLGIRR 544
Db 523 LKPVGTASAARRLGIRR 539
```

Search completed: June 1, 2004, 13:53:09
Job time : 33.7307 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 45.5302 Seconds
(without alignments)
3289.030 Million cell updates/sec

Title: US-09-926-799-2
Perfect score: 2819
Sequence: 1 MMASKDATSSVDGASGAGQ.....YQLKPVGTASSARGRLGLRR 530

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1980s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2819	100.0	530	4	AAB49701 Small rou
2	2793	99.1	530	2	AAR50972 Norwalk v
3	2793	99.1	530	7	ADC72176 Norwalk v
4	2713	96.2	530	2	AAR57091 Small rou
5	2093	74.2	544	4	AAB49703 Small rou
6	2029.5	72.0	546	4	AAB49702 Small rou
7	1938	68.7	545	4	AAB49700 Small rou
8	1232	43.7	540	4	AAB49706 Small rou
9	1209	42.9	548	4	AAB49705 Small rou
10	1200.5	42.6	535	4	AAB49707 Small rou
11	1188	42.1	548	5	AAB49702 Norwalk v
12	1181	41.9	542	4	AAB49708 Small rou
13	1148.5	40.7	539	4	AAB49704 Small rou
14	1131	40.1	550	4	AAB49709 Small rou
15	1005.5	35.7	541	4	AAB49710 Small rou
16	326	11.6	579	2	AAW08143 RHDV caps
17	244.5	8.7	668	4	AAB67462 Amino aci
18	240.5	8.5	668	2	AAR10686 Feline ca
19	240.5	8.5	668	4	AAE04304 Feline ca
20	229	8.5	623	4	AAB47043 Feline ca
21	238	8.4	669	4	AAB67461 Amino aci
22	237	8.4	547	4	AAW50108 Feline ca
23	237	8.4	671	4	AAW50107 Feline ca
24	236	8.4	623	4	AAB47044 Feline ca
25	225	8.0	622	4	AAB47045 Feline ca

26	136	4.8	40	5	AAU91273	Aau91273 Norwalk v
27	134.5	4.8	2164	1	AAP81045	Aap81045 Sequence
28	132.5	4.7	2164	1	AAP80131	Aap80131 Peptides
29	132	4.7	2206	2	AAR22210	Aar22210 True type
30	130	4.6	40	5	AAU91274	Aau91274 Norwalk v
31	127	4.5	517	2	AAW17975	Aaw17975 Coprinus
32	126.5	4.5	1049	6	ABU21976	Abu21976 Protein e
33	126	4.5	2209	1	AAP20037	Aap20037 Sequence
34	126	4.5	6310	6	ABU39869	Abu39869 Protein e
35	124.5	4.4	3892	6	ADA34216	Ada34216 Acinetoba
36	123.5	4.4	2183	3	AAB03533	Aab03533 Murine fa
37	123.5	4.4	2639	4	ABG15016	Abg15016 Novel hum
38	121.5	4.3	749	6	ABU27823	Abu27823 Protein e
39	119.5	4.2	3716	6	ABM15900	Abm15900 Mycobacte
40	119	4.2	1121	4	ABB59969	Abb59969 Drosophila
41	118.5	4.2	8360	6	AAE35499	Aae35499 Streptomy
42	118	4.2	918	2	AAW88422	Aaw88422 Chlamydia
43	117.5	4.2	1459	6	ABM15895	Abm15895 Mycobacte
44	117.5	4.2	1459	7	ADB80091	Adb80091 Mycobacte
45	116	4.1	2468	6	ABU38411	Abu38411 Protein e

ALIGNMENTS

RESULT 1
AAB49701
ID AAB49701 standard; protein; 530 AA.

AC AAB49701;
XX
DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 2.
XX
KW Small round structured virus; SRSV; food poisoning.
XX
OS Small round structured virus.

XX WO200079280-A1.

XX 28-DEC-2000.

XX 22-JUN-2000; 2000WO-JP004095.

XX 22-JUN-1999; 99JP-00175928.

XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
(DENK-) DENKA SEIKEN KK.

XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

XX WPI: 2001-080848/09.

XX N-PSDB; AAF29142.

XX Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.

XX Claim 1; Page 42-45; 84pp; Japanese.

XX This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

XX Sequence 530 AA;

Query Match 100.0%; Score 2819; DB 4; Length 530;
Best Local Similarity 100.0%; Pred. No. 1.5e-247;

Wed Jun 2 09:13:26 2004

us-09-926-799-2.rag

Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMASKDATSSVDGAGAGQVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIIINF 60
DB 1 MMASKDATSSVDGAGAGQVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIIINF 60

QY 61 VOAPQGEFTISPNNTPGGVLFDSLGLPHLPFLHLSQMYNGWGNMVRIMLAGNAFTA 120
DB 61 VOAPQGEFTISPNNTPGGVLFDSLGLPHLPFLHLSQMYNGWGNMVRIMLAGNAFTA 120

QY 121 GKIIVSCIPPGFSGHNLTTIAQATLFPFHVIADVRTLDPIEVLDPEDVRNVLPHNDRNQTM 180
DB 121 GKIIVSCIPPGFSGHNLTTIAQATLFPFHVIADVRTLDPIEVLDPEDVRNVLPHNDRNQTM 180

QY 181 RLVCMLYTLPLRTGGGSDSFVAGRVMTCPSPDFNLFVPTVEQKTRPFTLPNLPSS 240
DB 181 RLVCMLYTLPLRTGGGSDSFVAGRVMTCPSPDFNLFVPTVEQKTRPFTLPNLPSS 240

QY 241 LNSRAPLPISGMISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVIN 300
DB 241 LNSRAPLPISGMISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVIN 300

QY 301 LTLDGTPFHPFEGPAPIGFDPDLGGCDWHINMTQFGHSSQTYDVTTPDTFVPHLGSIQ 360
DB 301 LTLDGTPFHPFEGPAPIGFDPDLGGCDWHINMTQFGHSSQTYDVTTPDTFVPHLGSIQ 360

QY 361 ANGIGSNYIGVLSWSPSPHSPSGQVDLWKIPNYGSSITEATHLAPSVYPPGFGEVLVF 420
DB 361 ANGIGSNYIGVLSWSPSPHSPSGQVDLWKIPNYGSSITEATHLAPSVYPPGFGEVLVF 420

QY 421 FMSKIPGAYSLPCLLPQEIYISHLASEQAPTVEAALLHYVDPDGTGRTLGEFKAYPDGF 480
DB 421 FMSKIPGAYSLPCLLPQEIYISHLASEQAPTVEAALLHYVDPDGTGRTLGEFKAYPDGF 480

QY 481 LTCVPNGASSGPQQLPINGVVFVSWRSRFFYQLKPVGTASSARGRLGLRR 530
DB 481 LTCVPNGASSGPQQLPINGVVFVSWRSRFFYQLKPVGTASSARGRLGLRR 530

RESULT 2
AAR50972
ID AAR50972 standard; protein; 530 AA.
XX AC AAR50972;
XX DT 16-OCT-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 05-OCT-1994 (first entry)
XX DE Norwalk virus strain 8FIIa protein (encoded by ORF2).
XX KW Norwalk virus; pathogen; acute gastroenteritis; food poisoning;
XX KW seafood contamination; diagnostic assay; calcivirus; small round virus.
XX OS Norwalk virus; (strain 8FIIa).
XX PN **PO9405700-A2.**
XX PD 17-MAR-1994.
XX PF 07-SEP-1993; 93WO-US008447.
XX PR 07-SEP-1992; 92US-00941365.
XX PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX PI **Natson, DO, Bates MK, Jiang X, Graham DY;**
XX DR WPI; 1994-101125/12.
XX DR N-PSDB; AAQ56826.
XX PT DNA from Norwalk and related viruses - used for preparing prods. for use
PT in diagnostic assays, detection and vaccines for Norwalk and related

PT viruses.
XX Claim 14; Page 68-70; 156pp; English.
XX The Norwalk virus was isolated from stool samples from adult volunteers
CC infected with safety tested Norwalk virus strain 8FIIa. The coding
CC sequence is useful for the design of probes for use in diagnostic assays
CC for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct
CC FN field.) (Updated on 16-OCT-2003 to standardise OS field)
XX SQ Sequence 530 AA;

Query Match 99.1%; Score 2793; DB 2; Length 530;
Best Local Similarity 98.7%; Pred. No. 3.4e-245;
Matches 523; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MMASKDATSSVDGAGAGQVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIIINF 60
DB 1 MMASKDATSSVDGAGAGQVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIIINF 60

QY 61 VOAPQGEFTISPNNTPGGVLFDSLGLPHLPFLHLSQMYNGWGNMVRIMLAGNAFTA 120
DB 61 VOAPQGEFTISPNNTPGGVLFDSLGLPHLPFLHLSQMYNGWGNMVRIMLAGNAFTA 120

QY 121 GKIIVSCIPPGFSGHNLTTIAQATLFPFHVIADVRTLDPIEVLDPEDVRNVLPHNDRNQTM 180
DB 121 GKIIVSCIPPGFSGHNLTTIAQATLFPFHVIADVRTLDPIEVLDPEDVRNVLPHNDRNQTM 180

QY 181 RLVCMLYTLPLRTGGGSDSFVAGRVMTCPSPDFNLFVPTVEQKTRPFTLPNLPSS 240
DB 181 RLVCMLYTLPLRTGGGSDSFVAGRVMTCPSPDFNLFVPTVEQKTRPFTLPNLPSS 240

QY 241 LNSRAPLPISGMISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVIN 300
DB 241 LNSRAPLPISGMISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVIN 300

QY 301 LTLDGTPFHPFEGPAPIGFDPDLGGCDWHINMTQFGHSSQTYDVTTPDTFVPHLGSIQ 360
DB 301 LTLDGTPFHPFEGPAPIGFDPDLGGCDWHINMTQFGHSSQTYDVTTPDTFVPHLGSIQ 360

QY 361 ANGIGSNYIGVLSWSPSPHSPSGQVDLWKIPNYGSSITEATHLAPSVYPPGFGEVLVF 420
DB 361 ANGIGSNYIGVLSWSPSPHSPSGQVDLWKIPNYGSSITEATHLAPSVYPPGFGEVLVF 420

QY 421 FMSKIPGAYSLPCLLPQEIYISHLASEQAPTVEAALLHYVDPDGTGRTLGEFKAYPDGF 480
DB 421 FMSKIPGAYSLPCLLPQEIYISHLASEQAPTVEAALLHYVDPDGTGRTLGEFKAYPDGF 480

QY 481 LTCVPNGASSGPQQLPINGVVFVSWRSRFFYQLKPVGTASSARGRLGLRR 530
DB 481 LTCVPNGASSGPQQLPINGVVFVSWRSRFFYQLKPVGTASSARGRLGLRR 530

RESULT 3
ADC72176
ID ADC72176 standard; protein; 530 AA.
XX AC ADC72176;
XX DT 18-DEC-2003 (first entry)
XX DE Norwalk virus protein 2 amino acid sequence.
XX KW immune response; non-Norwalk virus agent; immunogen; Norwalk virus;
XX KW viral pathogen; acute gastroenteritis; virucidal; antiviral vaccine.
XX OS Norwalk virus.
XX PN **US6572862-B1.**
XX PD 03-JUN-2003.
XX PF 07-JUN-1995; 95US-00486049.

XX 08-NOV-1989; 89US-00433492.
 PR 27-APR-1990; 90US-00515993.
 PR 27-AUG-1990; 90US-00573509.
 PR 06-MAY-1991; 91US-00696454.
 XX (BAYU) BAYLOR COLLEGE MEDICINE.
 PA ~~Estes MK~~ Jiang X, Graham DY;
 PI WPI; 2003-776005/73.
 XX DR N-PSDB; ADC72174.
 XX DR N-PSDB; ADC72174.
 XX Inducing an immune response against non-Norwalk virus agents, comprises
 PT administering an immunogen recombinantly expressed from a cDNA from
 PT Norwalk virus.
 XX
 XX Example 4; SEQ ID NO 3; 45pp; English.
 XX
 CC This invention relates to a novel method of inducing an immune response
 CC in an individual against Norwalk virus and non-Norwalk virus agents, by
 CC orally or parenterally administering an immunogen recombinantly expressed
 CC or synthesised from a cDNA of Norwalk virus given in the specification.
 CC Norwalk virus is one of the most important viral pathogens, causing acute
 CC gastroenteritis. The invention may be used for the development of
 CC compounds with virucidal activity or an antiviral vaccine. The present
 CC sequence is the amino acid sequence of a protein encoded by the Norwalk
 CC virus genome of the invention.
 XX
 XX Sequence 530 AA;

Query Match 99.1%; Score 2793; DB 7; Length 530;
 Best Local Similarity 98.7%; Pred. No. 3.4e-245;
 Matches 523; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MMASKDATSSVDGASGAGQLVPEYNASDPLAMDPAVGSSTAVATAGQVNPIDPMIINF 60
 DB 1 MMASKDATSSVDGASGAGQLVPEYNASDPLAMDPAVGSSTAVATAGQVNPIDPMIINF 60
 QY 61 VOAPQGEFTISPNNTPGGVLFDLSLGPLHNPFLHLSQMYNGWGMVRIMLAGNAFTA 120
 DB 61 VOAPQGEFTISPNNTPGDVLFDLSLGPLHNPFLHLSQMYNGWGMVRIMLAGNAFTA 120
 QY 121 GKIIVSCIPPGGSHNLTIAQATLPHVIADVRTLDPIEVPLEDVRNVLFNHNDNRQQT 180
 DB 121 GKIIVSCIPPGGSHNLTIAQATLPHVIADVRTLDPIEVPLEDVRNVLFNHNDNRQQT 180
 QY 181 RLVCMLYTLPTLRTGGTGDSEFVAGRVMTCPSPDFNLFVLPPTVEQKTRPFTLPNPLSS 240
 DB 181 RLVCMLYTLPTLRTGGTGDSEFVAGRVMTCPSPDFNLFVLPPTVEQKTRPFTLPNPLSS 240
 QY 241 LNSRAPLPISGMGI SPDNVQSVQFQNGRCTLDGLRGVTPVSLSHVAKIRGTSNGTVN 300
 DB 241 LNSRAPLPISGMGI SPDNVQSVQFQNGRCTLDGLRGVTPVSLSHVAKIRGTSNGTVN 300
 QY 301 LTELDTGTPHPPEGPAPIGFDPDLGGCDWHINMTQFGHSSQTQYDVTDTFVPHLGSIQ 360
 DB 301 LTELDTGTPHPPEGPAPIGFDPDLGGCDWHINMTQFGHSSQTQYDVTDTFVPHLGSIQ 360
 QY 361 ANIGSGNVI GVLWSVPSHSGSQVDLWKIPNYGSSITEATHLAPSVYPPGFEVLVF 420
 DB 361 ANIGSGNVI GVLWSVPSHSGSQVDLWKIPNYGSSITEATHLAPSVYPPGFEVLVF 420
 QY 421 FMSKIPGAGYSILPCLLPOEYI SHLASEQAPTVGEAALLHYVDPTDTRGLTGFKAYPDGF 480
 DB 421 FMSKIPGAGYINLPCLLPOEYI SHLASEQAPTVGEAALLHYVDPTDTRGLTGFKAYPDGF 480
 QY 481 LTCVPNGASSGQOOLPINGVVFVSWSRFYQLKPVGTASSARGRLGLRR 530
 DB 481 LTCVPNGASSGQOOLPINGVVFVSWSRFYQLKPVGTASSARGRLGLRR 530

RESULT 4

AAR57091
 ID AAR57091 standard; protein; 530 AA.
 XX
 AC AAR57091;
 XX
 DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 05-OCT-1994 (first entry)
 XX
 DE Small round virus SRSV/KY/89 capsid protein.
 XX
 KW pathogen; acute gastroenteritis; food poisoning; seafood contamination;
 KW diagnostic assay; human calcivirus; small round virus; SRSV; KY89;
 KW Norwalk virus; capsid protein.
 XX
 OS Small round structured virus.
 XX
 PN WO9405700-A2.
 XX
 PD 17-MAR-1994.
 XX
 PF 07-SEP-1993; 93WO-US008447.
 XX
 PR 07-SEP-1992; 92US-00941365.
 XX
 PA (BAYU) BAYLOR COLLEGE MEDICINE.
 XX ~~Estes MK~~ Jiang X, Graham DY;
 XX WPI; 1994-101125/12.
 DR N-PSDB; AAO56832.
 XX
 CC DNA from Norwalk and related viruses - used for preparing prods. for use
 CC in diagnostic assays, detection and vaccines for Norwalk and related
 CC viruses.
 XX
 PS Example 7; Fig 13a; 156pp; English.
 XX
 CC The known sequence for Norwalk virus was used to obtain the sequence of
 CC other Norwalk-related viruses such as SRSV/KY/89, an agent from a stool
 CC from an outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide
 CC cDNA sequence includes part of the polymerase region and the capsid
 CC region of the genome; the deduced amino acid sequences are AAR57092 and
 CC AAR57091, respectively. Expression of fragments and derivs. of Norwalk-
 CC related viruses permits development of diagnostic assays to detect
 CC antibodies, antigens, viral genetic material or antivirals. (Updated on
 CC 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS
 CC field.)
 XX

Seq Sequence 530 AA;
 Query Match 96.2%; Score 2713; DB 2; Length 530;
 Best Local Similarity 96.8%; Pred. No. 6.6e-238;
 Matches 513; Conservative 2; Mismatches 15; Indels 0; Gaps 0;
 QY 1 MMASKDATSSVDGASGAGQLVPEYNASDPLAMDPAVGSSTAVATAGQVNPIDPMIINF 60
 DB 1 MMASKDATSSVDGASGAGQLVPEYNASDPLAMDPAVGSSTAVATAGQVNPIDPMIINF 60
 QY 61 VOAPQGEFTISPNNTPGGVLFDLSLGPLHNPFLHLSQMYNGWGMVRIMLAGNAFTA 120
 DB 61 VOAPQGEFTISPNNTPGDVLFDLSLGPLHNPFLHLSQMYNGWGMVRIMLAGNAFTA 120
 QY 121 GKIIVSCIPPGGSHNLTIAQATLPHVIADVRTLDPIEVPLEDVRNVLFNHNDNRQQT 180
 DB 121 GKIIVSCIPPGGSHNLTIAQATLPHVIADVRTLDPIEVPLEDVRNVLFNHNDNRQQT 180
 QY 181 RLVCMLYTLPTLRTGGTGDSEFVAGRVMTCPSPDFNLFVLPPTVEQKTRPFTLPNPLSS 240
 DB 181 RLVCMLYTLPTLRTGGTGDSEFVAGRVMTCPSPDFNLFVLPPTVEQKTRPFTLPNPLSS 240
 QY 241 LNSRAPLPISGMGI SPDNVQSVQFQNGRCTLDGLRGVTPVSLSHVAKIRGTSNGTVN 300

Db 241 LNSRAPLPISGMISPDNVQSVQFQNGRCITLDGLRVGTPPVSLSHVAKIRGTSNGTVIN 300
 QY 301 LTELDTTPHPPEGPAPIGFPDLGGCDHINNTQFCHSSQTOYDVTTPDTFVPHLGSIQ 360
 Db 301 LTELDTTPHPPEGPAPIGFPDLGGCDHINNTQFCHSSQTOYDVTTPDTFVPHLGSIQ 360
 QY 361 ANGISGNYIGVLSWVSPSHPSGQVDLWKIPNYGSSITEATHLAPSVYPPGFGEVLVF 420
 Db 361 ANGISGNYIGVLSWVSPSHPSGQVDLWKIPNYGSSITEATHLAPSVYPPGFGEVLVF 420
 QY 421 FMSKTPGCGAYSLPCLLPQYISHLASEQAPTVGEAALLHYVDPDTGRTLGFKAYPDGF 480
 Db 421 FMSKTPGCGAYSLPCLLPQYISHLASEQAPTVGEAALLHYVDPDTGRTLGFKAYPDGF 480
 QY 481 LTCVNGASSGQQLPVGFPVSVWVSRFYQLKPVGTASSARGRLGLRR 530
 Db 481 LTCVNGASSGQQLPVGFPVSVWVSRFYQLKPVGTASSARGRLGLRR 530
 RESULT 5
 AAB49703
 ID AAB49703 standard; protein; 544 AA.
 AC AAB49703;
 DT 04-APR-2001 (first entry)
 DE Small round structured virus protein SEQ ID 4.
 KW Small round structured virus; SRSV; food poisoning.
 OS Small round structured virus.
 PN WO200079280-A1.
 PD 28-DEC-2000.
 PF 22-JUN-2000; 2000WO-JP004095.
 PR 22-JUN-1999; 99JP-00175928.
 PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 DR WPI; 2001-080848/09.
 DR N-PSDB; AAF29144.
 PT Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 PS Claim 1; Page 47-49; 84pp; Japanese.
 XX This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX
 XX Sequence 544 AA;
 Query Match 74.2%; Score 2093; DB 4; Length 544;
 Best Local Similarity 72.7%; Pred. No. 2.1e-181;
 Matches 396; Conservative 53; Mismatches 80; Indels 16; Gaps 8;
 QY 1 MMASKDATSSVDGASGAGQLVBEVNASDPLAMDVAGSSSTAVATAGVNPIDPWIIINF 60
 Db 1 MMASKDATSPADGATGAGQLVBEVNTADPIPDPVAGSSSTALATAGVNLIDPWIIINF 60

QY 61 VQAPQGEFTTSPNNTPGGVLFDLGLGPHLAPFLHLHSQMYNGWGNVRIMLAGNAFTA 120
 Db 61 VQAPQGEFTTSPNNTPGGVLFDLGLGPHLAPFLHLHSQMYNGWGNVRIMLAGNAFTA 120
 QY 121 GKIVSCIPPGFGSHNUTIAQATLFPVHIADVRLTDIEVPLEDYRVNVLPHNNRNQOTM 180
 Db 121 GKIVSCIPPGFGSHNUTIAQATLFPVHIADVRLTDIEVPLEDYRVNVLPHNNRNQOTM 180
 QY 181 RLVCMLYTPRLTGGGTG--DSFVVAGVMTCPSPDENFLPLVPTVEKTRPFTLPNLP 238
 Db 180 RLLCMLYTPRLTGGGTGDSFVVAGVMTCPSPDENFLPLVPTVEKTRPFTLPNLP 239
 QY 239 SSLNSRAPLPISGMISPDNVQSVQFQNGRCITLDGLRVGTPPVSLSHVAKIRG--TSNG 296
 Db 240 KYLSNSRIPNPIEGMSLSPDQTQNVQFQNGRCITLDGLRVGTPPVSVLSQCKFRGRTISGQ 299
 QY 297 TVINLTLDGTPPHFPPEGPAPIGFPDLGGCDHINNTQFCHSSQTOYDVTTPDTFVPH 351
 Db 300 RVLNLTLDGTPPHFPPEGPAPIGFPDLGGCDHINNTQFCHSSQTOYDVTTPDTFVPH 359
 QY 352 FVPHLGSIT--QANGIGSGNYIGVLSWVSPSHPSGQVDLWKIPNYGSSITEATHLAPSV 409
 Db 360 FVPHLGSITLDENSSGGDIYIGTQWTSPPSDSGGANTFNFKIPDYGSSLAESQALAPAV 419
 QY 410 YPPGFGEVLVFFMSKIPGP---GAYSL-PCLLPQYISHLASEQAPTVGEAALLHYVDP 465
 Db 420 YPPGFGEVLVFFMSKIPGP---GAYSL-PCLLPQYISHLASEQAPTVGEAALLHYVDP 479
 QY 466 TGRITLGEFKAYPDGFLTCVNGASSGQQLPVGFPVSVWVSRFYQLKPVGTASSARGR 525
 Db 480 TNRNLGEFLYPPGGYLTCPVNSSSTGTPQQLPLDGVFVFAVSWVSRFYQLKPVGTAGPARGR 539
 QY 526 LGRLR 530
 Db 540 LGVRR 544
 RESULT 6
 AAB49702
 ID AAB49702 standard; protein; 546 AA.
 AC AAB49702;
 DT 04-APR-2001 (first entry)
 DE Small round structured virus protein SEQ ID 3.
 KW Small round structured virus; SRSV; food poisoning.
 OS Small round structured virus.
 PN WO200079280-A1.
 PD 28-DEC-2000.
 PF 22-JUN-2000; 2000WO-JP004095.
 PR 22-JUN-1999; 99JP-00175928.
 PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 DR WPI; 2001-080848/09.
 DR N-PSDB; AAF29143.
 PT Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 PS Claim 1; Page 45-47; 84pp; Japanese.
 XX

CC This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks

XX Sequence 546 AA;

Query Match 72.0%; Score 2029.5; DB 4; Length 546;
 Best Local Similarity 69.4%; Pred. No. 1.3e-175;
 Matches 379; Conservative 68; Mismatches 82; Indels 17; Gaps 6;

QY 1 MMWASKDATTSSVDGASGAGQLVPEVNASDPLAMPDVAGSSTAVATAGQVNPIDPMIINN 60
 DB 1 MMWASKDAPQSDAGSAGAGQLVPEVNTADPLMPVAGPTTAVATAGQVNMIDPMIINN 60
 QY 61 VQAPQGEFTISPNNTPGGVLDLSLGLPHLNPFLHLSQMYNGWGVNMRVIRIMLAGNAFTA 120
 DB 61 VQSPQGEFTISPNNTPGDILFDLQGLPHLNPFLHLSQMYNGWGVNMRVIRILLAGNAFSA 120
 QY 121 GKIIIVSCIPPGGSHNLTAQATLPHVIADVRTLDPIEVPLEDVNVLFNHNDRNQOQM 180
 DB 121 GKIIIVCCVPPGFTSSSLTTAQTLPFHVIADVRTLEPIEMPLEDVNVLVHTND-NQPTM 179
 QY 181 RLVCMLYTLPLRTGGGTG--DSFVAVAGVMTCPSPDFNLFVLPVPTVEQKTRPFTLPNPL 238
 DB 180 RLVCMLYTLPLRTGGSGNSDSFVAVAGVLTAPSSDFSFLVLPVPTVEQKTRPFTLPNPL 239
 QY 239 SLSNSRABLPISGMGISPDNVQSQFQNGRCTLDGRVLGTTVPVSLSHVAKIRGTSN--G 296
 DB 240 QTLNSRFPESLIQGMILSPDASQVQFQNGRCLIDQLGLTTPATSGQLFRVRGKINQGA 299
 QY 297 TWINLTGDTPEHPPEGPAPICGFDPLGGCDWHINNTQFGHSSQT-----QYDVTPTPT 351
 DB 300 RTUNLTFVQKGFMAFDSFAPVGFDFGKCDWHMRISKTPNNTSSGDPNRSVQVNTVOG 359
 QY 352 FVPHLGSIQANGI--GSGNYIGVLSWVSPSPHSPGSDQLWKIPNYGSSITEATHLAPS 408
 DB 360 FVPHLGSIQDFVFNHTPDYGTIEWISQSTPTPCTDINLEIPIDYGSSLSQAANLAPP 419
 QY 409 VTPPGGEVLVFMFKIPGFGAYS-----LPCLLPQBYISHLASEQAPTVGEAALLHYVDP 464
 DB 420 VTPPGGFEALVYFVSAPFGPNRNSAPNDVPCLLPQBYITHFVSEQAPTMDAALLHYVDP 479
 QY 465 DTGRTLGEFKAYPDGFLTCVPNGASSGPQOLPINGVVFVSVWSRYQLKPVGTASSARG 524
 DB 480 DTNRNLGEBFKLYPGGYLTCVPNGVGAGPQOLPLNGVFLFVSWVSRFYQLKPVGTASTARS 539
 QY 525 RLGLRR 530
 DB 540 RLGVRR 545

RESULT 7

AAB49700
 ID AAB49700 standard; protein; 545 AA.

XX AAB49700;

XX 04-APR-2001 (first entry)

XX Small round structured virus protein SEQ ID 1.

XX Small round structured virus; SRSV; food poisoning.

XX Small round structured virus.

XX WO200079280-A1.

XX 28-DEC-2000.

XX

PF 22-JUN-2000; 2000WO-JP004095.
 XX 22-JUN-1999; 99JP-00175928.
 PR
 XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.
 XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 PI WPI; 2001-080848/09.
 DR N-PSDB; AAF29141.
 XX
 PT Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX
 PS Claim 1; Page 40-42; 84pp; Japanese.
 XX
 CC This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks

XX Sequence 545 AA;

Query Match 68.7%; Score 1938; DB 4; Length 545;
 Best Local Similarity 67.0%; Pred. No. 2.7e-167;
 Matches 366; Conservative 63; Mismatches 99; Indels 18; Gaps 7;

QY 1 MMWASKDATTSSVDGASGAGQLVPEVNASDPLAMPDVAGSSTAVATAGQVNPIDPMIINN 60
 DB 1 MMWASKDAPTNDGTSGAGQLVPEANTAEPISEMPEVAGAAATAATAGQVNMIDPMIINN 60
 QY 61 VQAPQGEFTISPNNTPGGVLDLSLGLPHLNPFLHLSQMYNGWGVNMRVIRIMLAGNAFTA 120
 DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLHLSQMYNGWGVNMRVIRILLAGNAFTA 120
 QY 121 GKIIIVSCIPPGGSHNLTAQATLPHVIADVRTLDPIEVPLEDVNVLFNHNDRNQOQM 180
 DB 121 GKIIIVSCIPPGGSHNLTAQATLPHVIADVRTLDPIEVPLEDVNVLFNHNDRNQOQM 180
 QY 181 RLVCMLYTLPLRTGGGTG--DSFVAVAGVMTCPSPDFNLFVLPVPTVEQKTRPFTLPNPL 238
 DB 180 RLVCMLYTLPLRAGSSSGTDPFVIAGVLTCPSPDFSLVLPVNPVEQKTRPFTLPNPL 239
 QY 239 SLSNSRABLPISGMGISPDNVQSQFQNGRCTLDGRVLGTTVPVSLSHVAKIRGT-----S 294
 DB 240 NTLNSRVPESLTKSMVSRDHQMVQFQNGRVTLDGQLQGTTPTSASQLCKIRGSRVFHAN 299
 QY 295 NCTVINLTGDTPEHPPEGPAPICGFDPLGGCDWHIN--MTQFGHSSQTQYDVTPTPT 351
 DB 300 GNGYNLTGDSYHAFESPAPICGFDPLGGCDWHMEASPTTQFNTGDKVQINVKQBSA 359
 QY 352 FVPHLGSIQANGI--GSGNYIGVLSWVSPSPHSPGSDQLWKIPNYGSSITEATHLAP 407
 DB 360 FAPHLGTLQADGLSDVSVNTNMIKLGWVSPVSDGHRGDVDPWVPRYGSTLTEAQLAP 419
 QY 408 SVTPPGGEVLVFMFKIP--GPGAYSILPCLLPQBYISHLASEQAPTVGEAALLHYVDP 464
 DB 420 PIYPPGFEALVYFVMSDFPIAHGTINGLSVPCTIPQEFVTHFVNEQAPTRGEAALLHYLDP 479
 QY 465 DTGRTLGEFKAYPDGFLTCVPNGASSGPQOLPINGVVFVSVWSRYQLKPVGTASSARG 524
 DB 480 DTHRNLGEBFKLYPEGFMTCVPSNCGTGPQTUPINGVVFVSVWSRYQLKPVGTAGPA-C 538
 QY 525 RLGLRR 530
 DB 539 RLGIIR 544

[illegible]

343	-YD--VDTTPTDFVPHLSIQANGIGSNYIGV-----LSWVSPSPHSGSQVD	388
350	AADAVVATVSDKYTKPLGLVQ---IGTWNTNDVENQPTKFTPIGLNEVA-----NGHRFE	401
389	LWKIPNYGSSITEATHLAPSVYPPGFGGVLFVFMKIPGCGAYSLP---CLLPQEIYISHL	445
402	QWTLFRYSGALTLNNTLAPAVAPLFPFGRLLUFFRSYVLPKGGFGNPAIDCSVPQEWQHF	461
446	ASEQAPTGEAALLHYVPDPTGRTILGEPKAYPDGFLTCVPNGASGPOOLPINGVFVFS	505
462	YQESAPSLGDVALVRYVPDITGRVLFEAKLHKGGLTV--SSTSGTPVVVPANGYFKFDS	519
506	WYSRFRYQLKPVGTASSAR	523
520	WYNQFYSLAPMGTTGRR	537
QY	RESULT 9	
DB	AAB49705	
DB	ID	AAB49705 standard; protein; 548 AA.
QY	XX	AAB49705;
DB	XX	04-APR-2001 (first entry)
QY	XX	Small round structured virus protein SEQ ID 6.
DB	XX	Small round structured virus; SRSV; food poisoning.
QY	XX	Small round structured virus.
DB	XX	WO200079280-A1.
QY	XX	28-DEC-2000.
DB	XX	22-JUN-2000; 2000WO-JP004095.
QY	XX	22-JUN-1999; 99JP-00175928.
DB	XX	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
QY	XX	(DENK-) DENKA SEIKEN KK.
DB	XX	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
QY	XX	WPI; 2001-080848/09.
DB	XX	N-PSDB; AAF25146.
QY	XX	Kit for the detection and typing of small round-structured virus (SRSV)
DB	XX	strains for investigation of food poisoning outbreaks, contains
QY	XX	antibodies.
DB	XX	Claim 1; Page 52-54; 84pp; Japanese.
QY	XX	This invention relates to a kit for the detection and typing of small
DB	XX	round structured virus (SRSV) strains. The kit contains antibodies
QY	XX	directed against peptides represented in sequences AAB49700 - AAB49710,
DB	XX	which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
QY	XX	AAF20151 represent cDNA encoding the strain specific proteins. The kit is
DB	XX	used for detecting and typing strains of SRSV in order to prevent the
QY	XX	spread of infection and to examine the epidemiology of outbreaks
DB	XX	Sequence 548 AA;

Db 57 VQAPGGEFTVSRNPSGCVLLLELGPENIPYLAHLMRYNGYAGGFEVQVVLGNAFTA 116
Qy 121 GKIIIVSCIPPGGSHNLIAQATLPHVIADVRLDPIEVPLEDVNRNVLPHNNDRNQOTM 180
Db 117 GKIIIPAAIPNPIDNLNSAAQITMCPHVIVDRQLEFVNLPMDFVRNFFHYNQGSDSL 176
Qy 181 RLVCMLYTLPLRTGGTGDSEFVAGRVMTCPSPDFNPLFLVPTVQKTRPPTLPLPLSS 240
Db 177 RLIAMLYTLPLRANSGDDVFTVSCRVLTRPSDFSNFLVPTVBSKTPFTLPLITISE 236
Qy 241 LNSRAPLPISGMGLSPDNVQSFQNGRCTLDGRLVGTTPVSLSHVAKIRG-----TSN 295
Db 237 MNSRFPVPIESLHTSPTENIVVQNGRVTLDELGMGTQLPLSQICAFRGVLRSTRS 296
Qy 296 GT-----VINLTDLGTTPHPPEG-PAPIGFDDLGCDWHNNWTFQGHSS 339
Db 297 ASDQADATPRLFNYYVHVLQNLNGTPYDPAEDIPGLGTDFRG-----KVFQVAS 349
Qy 340 QTOYD-----VDTTPTDTPVPHLGSIOAN-----GIGSGNYIGVL 373
Db 350 QRLNLSSTRAHRAKVDTTAGRTPKLGSLEISTDSDDFQDQNPQTKFTPVGIGVDN----- 404
Qy 374 SVNSPPSPSGSQVDLWKIPNYGSSITEATHLAPSVYPPGFEVLVFPFNSKIPGPCAYS- 432
Db 405 -----EAEFQWLSUDYSGOFTHMNLAPAVAPNFPGEQLLFFRSQLPSSGGRSN 454
Qy 433 --LPCLLPOEYISHLASEQAPTVGEAALHYVDPTDGTGLGFEKAYPDGFLTCVPNGASS 490
Db 455 GVLDCLVQPEWQHVFQESAPAQTOVALVRYVNPDTGKVLFEAKLHKLGFMTIANGDS- 513
Qy 491 GQQQLPINGVVFVSVSVSFYOLKPVGTASSAR 523
Db 514 -FITVPPNGYFRFESWVNFYTLAPMGTGNRR 545

RESULT 10
ID AAB49707
XX AAB49707 standard; protein; 535 AA.
AC AAB49707;
XX 04-APR-2001 (first entry)
DT Small round structured virus protein SEQ ID 8.
DE Small round structured virus; SRSV; food poisoning.
KW Small round structured virus.
OS Small round structured virus.
XX WO200079280-A1.
XX 28-DEC-2000.
XX 22-JUN-2000; 2000WO-JP004095.
XX 22-JUN-1999; 99JP-00175928.
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI: 2001-080848/09.
DR N-PSDB; AAF29148.
XX Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX Claim 1; Page 57-59; 84pp; Japanese.
XX This invention relates to a kit for the detection and typing of small

CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX Sequence 535 AA;
SQ Query Match 42.6%; Score 1200.5; DB 4; Length 535;
Best Local Similarity 46.4%; Pred. No. 4.1e-100;
Matches 258; Conservative 76; Mismatches 165; Indels 57; Gaps 12;
Qy 1 MMASKDATTSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGVNPIDPMIINF 60
Db 1 MMASNDAAAPSNDGAAG---LVPEAN-NETMALEPVAGASIAAPLTGQNNIIDPMIRLNF 56
Qy 61 VOAPGEBETISNNTPGGVLFDLISGPHLNPPLHLSQMYAGVGNMVRIMLAGNAFTA 120
Db 57 VOAPNGEFTVSRNPSGCVLLLELGPENIPYLAHLMRYNGYAGGFEVQVVLGNAFTA 116
Qy 121 GKIIIVSCIPPGGSHNLIAQATLPHVIADVRLDPIEVPLEDVNRNVLPHNNDRNQOTM 180
Db 117 GKLVFAAAPPHPPLENISPGQITMFPVLIIDVRLTLEPVLLPLPDVRRNFFHYNQONEPRM 176
Qy 181 RLVCMLYTLPLRTGGTGDSEFVAGRVMTCPSPDFNPLFLVPTVQKTRPPTLPLPLSS 240
Db 177 RLVCMLYTLPLRANSGDDVFTVSCRVLTRPSDFNPLVPTVBSKTPFTLPLITIGE 236
Qy 241 LNSRAPLPISGMGLSPDNVQSFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTV-- 298
Db 237 LTNSRFPVPIDELYTSPNESLVQPPQNGRCALDGELOQTQLLPTAICSFGRIRINQKVS 296
Qy 299 -----INLTDLGTTPHPF-EGPAPIGFDDLGCDWHNNWTFQGHSSQTOYD----- 344
Db 297 ENHVMNQVTININGTTPDTPGVPAPLGTTPFSG-----KLFGLVSLQRDHNDACRSHD 349
Qy 345 --VDTTPTDTPVPHLGSIOANGIGSGNYIGVLSWSVSPSSH-----PSG-----SQVDLW 390
Db 350 AVIATNSAKFTPKLGAIQ-----IG--TWBDDVHINQPTKFTPVGLPENEGFNQW 398
Qy 391 KIPNYGSSITEATHLAPSVYPPGFEVLVFPFNSKIPGPCAYSPL---CLLPOEYISHLAS 447
Db 399 TLPNYSGALTLNMLGAPFVAPTFFGGEQLLFRSHIPLKGGVADPVIDCLLPOEWIQLYQ 458
Qy 448 EOAPTVGEAALHYVDPTDGTGLGFEKAYPDGFLTCVPNGASSGQQPLPINGVVFVSVW 507
Db 459 ESAPSQSDVALIRFTNPTDGTGRVLFPAKLRSGYITVANTG--SRFIVVPANGYFRFDTWV 516
Qy 508 SRFYOLKEVGTASSAR 523
Db 517 NQFVSLAPMGTGNRR 532
RESULT 11
ID AAU91272
XX AAU91272 standard; protein; 548 AA.
AC AAU91272;
XX 18-JUN-2002 (first entry)
DT Norwalk virus associated polynucleotide #1.
DE Norwalk virus associated polynucleotide #1.
KW Nowalk virus; monoclonal antibody; geno group I; geno group II;
KW immunological detection; food; viral infection.
OS Norwalk virus.
XX JP2002020399-A.
XX 23-JAN-2002.

PF 10-JUL-2000; 2000JP-00208151.
FR 10-JUL-2000; 2000JP-00208151.
XX (OSAP) OSAKA PREFECTURE.
PA (NISE-) NIPPON SEIBUTSU KAGAKU CENT KK.
PA (IATR) IATRON LAB INC.
XX WPI; 2002-287412/33.
XX A monoclonal antibody useful in the immunological detection and diagnosis
PT of Norwalk virus infection.
XX Disclosure; Page 12-13; 24pp; Japanese.
XX The invention describes a monoclonal antibody recognising Norwalk virus,
CC a capsid protein of Norwalk virus, or a common antigen epitope on the
CC capsid protein molecule of geno group I and geno group II. The antibody
CC is useful for immunological detection and quantitative analysis of
CC Norwalk virus in foods and the serum of infected patients. This sequence
CC represents a Norwalk virus associated protein described in the invention
XX
XX Sequence 548 AA;
Query Match 42.1%; Score 1188; DB 5; Length 548;
Best Local Similarity 45.3%; Pred. No. 5.8e-99;
Matches 258; Conservative 70; Mismatches 164; Indels 78; Gaps 11;
QY 1 MMWASKDATSSVDCASGAGOLVPEVNASDPLAMDVPVAGSSTAVATAGQVNPIDPWIINN 60
DB 1 MKWASRAAFNSDGAAG---LVPEIN-NEAALDPPVAGAAIAAPLTGQQNIIDPWIINN 56
QY 61 VQAPQGEFTTSPNNTPGGVLFDSLGLPHLNPFLHLSOMYNGVGNVRIMLAGNAFTA 120
DB 57 VQAPCGEFTVSPRNSPGCEVLLNLELGPENPYLAHLARMYNGVAGGFEVQVVLGNAFTA 116
QY 121 GKTIIVSCIPGFGSHNLTIAQATLFPHVIAVRLTDPIEVPLEDVRNVLPHNNDRNQTM 180
DB 117 GKTIIFAAIPNFPIDNLSAAQITMCPHVIIVDVRQLEPNLPMVVRNPFHYNQGSRL 176
QY 181 RLVCMLYTLRTGGTGDSFVVAGRVMTCPSPDNFLVPPVTEQKTRPTLNLPLSS 240
DB 177 RLIAMLYTLPLRANSGDDVFTVSCVLTFRSPDFSENFVLPPTVESKTKPTLPLTISE 236
QY 241 LNSRAPLPISGMGISPDNVSQVQFQNGRCITLDELGLVGTTPVSLSHVAKIRG-----TSN 295
DB 237 MNSRFPVPIESLHTSPTENIIVQCONGRVTLDELGMGTTLQLPSRICAFRGVLTSTR 296
QY 296 GT-----VINLTDLGTGTFPHFEG-PAPIGPDLGGCDWHNNMTQFGHSS 339
DB 297 ASDQADIATPLFNYYWHVQJDNLNGTPTDPAEDIPGLGTPDRG-----KVFGVAS 349
QY 340 QTQYD-----VDTTPTDFVPHLGSIQAN-----GIGSGNYIGVL 373
DB 350 QENPSTTTHAEAKVDYTAGRTPKLGSLEISTESSDQDQNTQTRFTPVGIGVDN----- 404
QY 374 SWVSPSPHSGQVDLWKIPNYGSSITEATHLASVYPGGEVLVFMKIPGPGAYS- 432
DB 405 -----EADFOQSLPDYSGQFTNNMLAPAVAFNPFGEQLLFRSOLPSSGGRSN 454
QY 433 --LPCLLPQEYISHLASEQAPTVEAALLHVDPDTRTGLGEKAYPDGFLTCVPNGASS 490
DB 455 GVLDCLVPOEWQHFYQESAPAQQTQVALVRYVNPDTGRVLPFAKLHLKLGFTIAKNGDS- 513
QY 491 GPQQLPINGVFFVSVWSRFPYOLKPVGTAS 520
DB 514 -FITVPPNGYFRFESWVNEFYTLAPMGTCN 542

RESULT 12
ID AAB49708
XX AAB49708 standard; protein; 542 AA.

AAC49708;
04-APR-2001 (first entry)
Small round structured virus protein SEQ ID 9.
Small round structured virus; SRSV; food poisoning.
Small round structured virus.
W0200079280-A1.
28-DEC-2000.
22-JUN-2000; 2000WO-JP004095.
22-JUN-1999; 99JP-00175928.
(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
(DENK-) DENKA SEIKEN KK.
Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
WPI; 2001-080848/09.
N-PSDB; AAF29149.
Kit for the detection and typing of small round-structured virus (SRSV)
strains for investigation of food poisoning outbreaks, contains
antibodies.
Claim 1; Page 59-61; 84pp; Japanese.
This invention relates to a kit for the detection and typing of small
round structured virus (SRSV) strains. The kit contains antibodies
directed against peptides represented in sequences AAB49700 - AAB49710,
which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
AAF20151 represent cDNA encoding the strain specific proteins. The kit is
used for detecting and typing strains of SRSV in order to prevent the
spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 542 AA;
Query Match 41.9%; Score 1181; DB 4; Length 542;
Best Local Similarity 46.7%; Pred. No. 2.5e-98;
Matches 259; Conservative 82; Mismatches 166; Indels 48; Gaps 14;
QY 1 MMWASKDATSSVDCASGAGOLVPE-VNASDPLAMDVPVAGSSTAVATAGQVNPIDPWIINN 59
DB 1 MKWASNDAAPSNDGAA---SLVPEGINETMPL--EPVAGASIAAPVAGQTNIDPWIINN 55
QY 60 FVQAPQGEFTTSPNNTPGGVLFDSLGLPHLNPFLHLSOMYNGVGNVRIMLAGNAFT 119
DB 56 FVQAPNGEFTVSPRNSPGCEILLNLELGPENPYLAHLARMYNGVAGGVEVQVVLGNAFT 115
QY 120 AGKIIIVSCIPGFGSHNLTIAQATLFPHVIAVRLTDPIEVPLEDVRNVLPHNNDRNQOT 179
DB 116 AGKILFAATPPNPLVDNISPAQITWDLPHLIVDVRILEPIMTLPDVRNVFYHFNQPPQR 175
QY 180 MRLVCMYTLPLRTGGTGDSFVVAGRVMTCPSPDNFLVPPVTEQKTRPTLNLPLS 239
DB 176 MRLVAMLYTLPLRANSGDDVFTVSCVLTFRTPDEFEFVLPVPSVESKTKPTLPLTIS 235
QY 240 SLNSRAPLPISGMGISPDNVSQVQFQNGRCITLDELGLVGTTPVSLSHVAKIRG---TSNG 296
DB 236 ELTNSRFPPIEQLYTAPNETNVVQCONGRCTLDGELQGTTLQLSSAVCFLOGRTVADNG 295
QY 297 -----TVINLTDLGTGTFPHF-EGPAPIGFPDLGGCDW-----HNNMTQFGHSSQTQYD 344
DB 296 DNWDQNLQLTYPNAGSYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGY 355
QY 345 VDTTPTDFVPHLGSIQANGIGSGNYIGVLSWSPSSH-----PGSGOVD-----LWKIPN 394
DB 356 ISITSGKFTPKIGSI---GLHS-----ITEHVHPNQOSRFTPVGVAVDENTFPQQWVLPH 407

QY 395 YGSSITEATHLAPSVYPPGFGVLFVFMKIP-----GPGAYSILPCLLPQBYISHLASE 448
 DB 408 YAGSLALNTLAPAVAPTTPGQQLFFRSRVCVQGLQDQAF-IDCLLPQEWVNFYQE 466
 QY 449 QAPTVEALHYVDPDGTGRTLGEFKAYPDGFLTCVPGNCASSGPOOLPINGVVFVSVWS 508
 DB 467 AAPSQADVALIRYNPDGTGRTLFEAKLHRSFGFITVSHTGAY--PLVVPNGHFRFDSWVN 524
 QY 509 REYQLKPVGTASSAR 523
 DB 525 QYSLAPMGTGNRR 539

RESULT 13
 AAB49704
 ID AAB49704 standard; protein; 539 AA.
 XX
 AC AAB49704;
 XX
 DT 04-APR-2001 (first entry)
 XX
 DE Small round structured virus protein SEQ ID 5.
 XX
 KW Small round structured virus; SRSV; food poisoning.
 XX
 OS Small round structured virus.
 XX
 PN WO200079280-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 22-JUN-2000; 2000WO-JP004095.
 XX
 PR 22-JUN-1999; 99JP-00175928.
 XX
 PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.
 XX
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 XX
 DR WPI; 2001-080848/09.
 DR N-PSDB; AAF29145.
 XX
 PT Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX
 PS Claim 1; Page 50-52; 84pp; Japanese.
 XX
 CC This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX
 SQ Sequence 539 AA;

Query Match 40.7%; Score 1148.5; DB 4; Length 539;
 Best Local Similarity 44.8%; Pred. No. 2.3e-95;
 Matches 247; Conservative 85; Mismatches 182; Indels 37; Gaps 14;

QY 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVAGSSTAVATAGQVNPIDPWNNF 60
 DB 1 MMASNDANPS-DGST--ANLVEYN-NEVMALEFVVGAAIAAPVAGQONVIDPWIRNF 56
 QY 61 VOAPQGEFTISPNTPGGVLPDLSPHLNPPLLHLSQWYNGWGNMVRIMLAGNAFTA 120
 DB 57 VOAPGGETVSPRNAPGILMSAPLCPDLNPNLHMYNGYAGGFEVQVILAGNAFTA 116
 QY 121 GKIIIVSCIPPGFGSHNLTTAQTLPFPHVIADVRLDPIEVPLEDVRLNLFHNDRNQTM 180

DB 117 GKIIFAAAPPNFPTEGLSPSQVTMPFPHIIVDRQLPVLILPLPVRNNFYHYNQNSDSTI 176
 QY 181 RLVCMLYTPRLTGGTGDGDFVAGRVMTCPSPDFNLFVLPPTVQKTRPFTLPNLPSS 240
 DB 177 KLIAMLYTPLRANNAGDDVFTVSCRVLTRPSDFDFELVPTVESRTKPTVPILTVEE 236
 QY 241 LSNRSAPLIPISGMGISPDNVQSQNGRCCTLDGRLVGTTPVLSLHVAKIRGT-----S 294
 DB 237 MSNSRFPILPEKLYTGPSSAFVQVQNGRCCTTDGVLGTTQLLSAVNICTFRGDVTTHIAGS 296
 QY 295 NCTVINLTDLGTPPHPEG-PAPIGFPD-LGGCDWHINMT--QFGHSSQQTQYDVTTPD 350
 DB 297 HDYTNWLASQWNSNYDPTTEIPAPLGTDFVKGIOGMLTQTTREDGSTRAHKATVSTGVS 356
 QY 351 TFPVHLGSIQ-----ANGIGSGNY-----IGVLSWVSPSPSGSQVDLWKIPNYGSSIT 400
 DB 357 HFTPKLGSVQYTTDNTNDFQTGNTKFTPVGVQI--DGNNHQNEPQ--QWVLPNYSGRGT 412
 QY 401 EATHLAPSVYPPGFGVLFVFMKIPGPGAY---SLPCLLPQBYISHLASEQAPTVGEAA 457
 DB 413 HNVHLAPAVAPTTPGQQLFFRSRVCVPGNCASSGPOOLPINGVVFVSVWSRVOLKP 472
 QY 458 LLHYVDPDGTGRTLGEFKAYPDGFLTCVPGNCASSGPOOL--PINGVVFVSVWSRVOLKP 515
 DB 473 LLRFVNPDTGRVLFECGLHKSGYTV----AHTGPHDLVIPNGYFRFDSWVNPQFYTLAP 528
 QY 516 VGTASSARGRL 526
 DB 529 MONGAGRREAL 539

RESULT 14
 AAB49709
 ID AAB49709 standard; protein; 550 AA.
 XX
 AC AAB49709;
 XX
 DT 04-APR-2001 (first entry)
 XX
 DE Small round structured virus protein SEQ ID 10.
 XX
 KW Small round structured virus; SRSV; food poisoning.
 XX
 OS Small round structured virus.
 XX
 PN WO200079280-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 22-JUN-2000; 2000WO-JP004095.
 XX
 PR 22-JUN-1999; 99JP-00175928.
 XX
 PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.
 XX
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 XX
 DR WPI; 2001-080848/09.
 DR N-PSDB; AAF29150.
 XX
 PT Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX
 PS Claim 1; Page 62-64; 84pp; Japanese.
 XX
 CC This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is

CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX
SQ Sequence 550 AA;

Query Match 40.1%; Score 1131; DB 4; Length 550;
Best Local Similarity 43.2%; Pred. No. 9.2e-94; Indels 78; Gaps 12;
Matches 248; Conservative 77; Mismatches 171; Indels 78; Gaps 12;

QY 1 MMASDKATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIINNF 60
Db 1 MKMASNDAAAPSSDGAAG--LVPEAN-DEWALEPVGASIAAPVVGQNIIDPWIRENF 56

QY 61 VQAPQGEFTISPNNTGGVLFDSLGPHNLPFLHLSQMYGWNVGMVRLMAGNAFTA 120
Db 57 VQAPQGEFTVSPRNSPGEMLNLELGPENLPYLSHLSMYNGYAGMGQVQVVLGNAFTA 116

QY 121 GKIIIVCIPGSGSHNLTAQATLPHVIADVTLDPIEVLDPEDVRNVLFNHNDROOQM 180
Db 117 GKIIIFAAVPPHPPFVENISAQAQITMCPHVIIDVRQLEPVLPLPDINRFFHYNQENTPRM 176

QY 181 RLVCMLYTLRTGGTGDGFVAGRVMTCPSPDFNLFVLPVPTVQKTRPFTLPLNPLSS 240
Db 177 RLVCMLYTLRANSGB-DVFTVSCRVLTRPAPDFEFTFLVPTVESKTKPFTLPLTIGE 235

QY 241 LNSRAPLPISGMIGISPDNVQSFQNGRCTLDGRVLTGTPVSLSHVAKIRGT----- 293
Db 236 LNSRFPAAIDMLYTDPNESIVVQPNQGRCTLDGTLQGTQVPTQICAFRGLISQATAR 295

QY 294 -----SNGTVINLTDLGTPFHPFEG-PAPIGFPDLGGCDWHINMTQFCHSSQ 340
Db 296 AADSTDSQARHNLPHVQVKNLDGTQYDFTDDIPAVLGAIDFKG-----TVFGVASQ 348

QY 341 T-----QYDVTPTDPTFVPHLGSIQANGISGNY-----IGV 372
Db 349 RDVSGQEQGHYATRAHEAHIDTDPKYAPKLTILIKS-GSDDFNTNQPRTPTVGMGD 407

QY 373 LSWVSPSPHSGQVDLWKIPNYGSSITEATHLAPSVYPPGCEVLVFFMSKIPGAYS 432
Db 408 NNRQ-----WELPDVSGRLTNLNNLAPAVSPSPGGERILFFRSIVPSAGYG 455

QY 433 ---LPCLLPQBYISHLASEQAPTVGEAALLHYDDPTGRTLGFBKAYPDGFLTCVPNGAS 489
Db 456 SGYIDCLIPQEWQHFQEAAPSQSAVALVRYNPDTRNIFEAKLHREGFLTVANCG-- 513

QY 490 SGPOOLPINGVVFVSVWSRPFYQLKPVGTASSAR 523
Db 514 NNPIVWPPNGYFRFEAWGNQFTYLLAPMGSGQGR 547

RESULT 15
AAB49710
ID AAB49710 standard; protein; 541 AA.
XX
AC AAB49710;
XX
DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 11.
XX
KW Small round structured virus; SRSV; food poisoning.
XX
OS Small round structured virus.
XX
PN WO200079280-A1.
XX
PD 28-DEC-2000.
XX
PP 22-JUN-2000; 2000WO-JP004095.
XX
PR 22-JUN-1999; 99JP-00175928.
XX
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

(DENK-) DENKA SEIKEN KK.
Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
WPI; 2001-080848/09.
N-PSDB; AAF29151.
Kit for the detection and typing of small round-structured virus (SRSV)
strains for investigation of food poisoning outbreaks, contains
antibodies.
Claim 1; Page 64-66; 84pp; Japanese.
This invention relates to a kit for the detection and typing of small
round structured virus (SRSV) strains. The kit contains antibodies
directed against peptides represented in sequences AAB49700 - AAB49710,
which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
AAF20151 represent cDNA encoding the strain specific proteins. The kit is
used for detecting and typing strains of SRSV in order to prevent the
spread of infection and to examine the epidemiology of outbreaks

Sequence 541 AA;
Query Match 35.7%; Score 1005.5; DB 4; Length 541;
Best Local Similarity 41.0%; Pred. No. 2.4e-82;
Matches 236; Conservative 70; Mismatches 179; Indels 91; Gaps 15;

QY 1 MMASDKATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIINNF 60
Db 1 MKMASNDAAAPSSDGAAG--LVPEIN-NEVMPLEPVAGASLATPVVGQNIIDPWIRNPF 56

QY 61 VQAPQGEFTISPNNTGGVLFDSLGPHNLPFLHLSQMYGWNVGMVRLMAGNAFTA 120
Db 57 VQAPAGEFTVSPRNSPGEIILDLGDLNLPYLAHLMYNGHAGGNEVQIVLAGNAFTA 116

QY 121 GKIIIVCIPGSGSHNLTAQATLPHVIADVTLDPIEVLDPEDVRNVLFNHNDROOQM 180
Db 117 GKIIIFAAVPPHPPFVENISAQAQITMCPHVIIDVRQLEPVLPLPDINRFFHYNQNDPKL 176

QY 181 RLVCMLYTLRTGGTGDGFVAGRVMTCPSPDFNLFVLPVPTVQKTRPFTLPLNPLSS 240
Db 177 RLVCMLYTLRANSGB-DVFTVSCRVLTRPAPDFEFTFLVPTVESKTKPFTLPLTIGE 236

QY 241 LNSRAPLPISGMIGISPDNVQSFQNGRCTLDGRVLTGTPVSLSHVAKIRGT-----TSNGT 297
Db 237 MTNSRFPVVDVMTARNENQVQPNQGRCTLDGTLQGTQVPTQICAFRGLISQATAR 296

QY 298 V-----INLTDLGTPFHPFEG-PAPIGFPDLGGCDWHINMTQFCHSSQYD----- 344
Db 297 VASVYRMDMEITNDGTDPIDTDPGPIGSPDFQ-----ILFGVASORNKNEQNAPAT 349

QY 345 -----VDTTPTDPTFVPHLGSIQANGISGNYI---GVLSWVSP----- 378
Db 350 RAHEAIINTGDLHLCPISS-----SEIVLTSNPIRLCTNPQPLPOSGLRGTILIRSD 402

QY 379 -----PSHPSGQVDLWKIPNYGSSITEATHLAPSVYPPGCEVLVFFMSKIPG 427
Db 403 NGCHDMVGTSTPTTWPQ--QMRCSRSGNSCCSGHRYP--VPVMMNRVTMVLSHKSG 458

QY 428 PGAYSILPCLLPQBYISHLASEQAPTVGEAALLHYDDPTGRTLGFBKAYPDGFLTCVPNG 487
Db 459 FSTSTRK--LPQNLN-----RWPLIRFINPDTGRVLFEARLHKQGFITVAHTG 504

QY 488 ASSGPOOLPINGVVFVSVWSRPFYQLKPVGTASSAR 523
Db 505 --DNPIVWPPNGYFRFEAWGNQFTYLLAPMGSGQGR 538

Search completed: June 1, 2004, 13:45:46
Job time : 47.5302 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 12.6275 Seconds
(without alignments)
2166.837 Million cell updates/sec

Title: US-09-926-799-2
Perfect score: 2819
Sequence: 1 MMASKDATSSVDGASGAGQ.....YOLKPVGTASSAGRLGLRR 530

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA*
1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS-COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2793	99.1	530	US-08-486-049-3	Sequence 3, Appli
2	244.5	8.7	668	US-09-617-594A-4	Sequence 4, Appli
3	239	8.5	623	US-09-590-020-2	Sequence 2, Appli
4	239	8.5	626	US-09-590-020-7	Sequence 7, Appli
5	238	8.4	669	US-09-617-594A-2	Sequence 2, Appli
6	236	8.4	623	US-09-590-020-4	Sequence 4, Appli
7	225	8.0	622	US-09-590-020-6	Sequence 6, Appli
8	129	4.6	2206	US-07-852-260-2	Sequence 2, Appli
9	129	4.6	2206	US-08-461-503-2	Sequence 2, Appli
10	129	4.6	2206	US-08-465-250-2	Sequence 2, Appli
11	127	4.5	517	US-08-689-421-33	Sequence 33, Appli
12	127	4.5	517	US-09-389-528-33	Sequence 33, Appli
13	127	4.5	517	US-09-181-827A-33	Sequence 33, Appli
14	124.5	4.4	3892	US-09-328-352-5503	Sequence 5503, Ap
15	123.5	4.4	2183	US-08-746-111-5	Sequence 5, Appli
16	107.5	3.8	1048	US-09-171-699-10	Sequence 10, Appli
17	106	3.8	20	US-08-973-961-5	Sequence 5, Appli
18	106	3.8	20	US-09-705-621-5	Sequence 5, Appli
19	106	3.8	1042	US-08-928-361B-11	Sequence 11, Appli
20	106	3.8	1042	US-09-588-995A-11	Sequence 11, Appli
21	106	3.8	1837	US-08-928-361B-5	Sequence 5, Appli
22	106	3.8	1837	US-09-588-995A-5	Sequence 5, Appli
23	105	3.7	1043	US-08-928-361B-30	Sequence 30, Appli
24	105	3.7	1721	US-08-700-651-5	Sequence 5, Appli
25	105	3.7	1721	US-08-928-361B-6	Sequence 6, Appli
26	105	3.7	1721	US-09-588-995A-6	Sequence 6, Appli
27	105	3.7	2736	US-09-252-991A-30227	Sequence 30227, A

28	102.5	3.6	770	4	US-09-489-039A-7872	Sequence 7872, Ap
29	102	3.6	539	2	US-08-808-931-16	Sequence 16, Appl
30	102	3.6	539	3	US-08-808-323-16	Sequence 16, Appl
31	102	3.6	539	3	US-09-050-603A-16	Sequence 16, Appl
32	102	3.6	539	3	US-09-102-420B-16	Sequence 16, Appl
33	102	3.6	539	4	US-09-497-698-16	Sequence 16, Appl
34	101.5	3.6	1234	2	US-08-317-310A-15	Sequence 15, Appl
35	101.5	3.6	1234	5	PCT-US95-13041-15	Sequence 15, Appl
36	100.5	3.6	2318	3	US-09-091-219-24	Sequence 24, Appl
37	100.5	3.6	2318	4	US-09-660-541-24	Sequence 24, Appl
38	99	3.5	512	1	US-08-462-484-6	Sequence 6, Appli
39	99	3.5	512	1	US-08-441-147-6	Sequence 6, Appli
40	99	3.5	512	5	PCT-US95-07536-6	Sequence 6, Appli
41	98.5	3.5	803	4	US-09-543-681A-5399	Sequence 5399, Ap
42	97	3.4	469	4	US-09-479-645A-6	Sequence 6, Appli
43	97	3.4	469	4	US-09-479-645A-8	Sequence 8, Appli
44	97	3.4	2442	4	US-09-514-247A-10	Sequence 10, Appli
45	96	3.4	2556	1	US-08-185-432-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-08-486-049-3
; Sequence 3, Application_US/08486049
; Patent No. 6572862
; GENERAL INFORMATION:
; APPLICANT: Bates, Mary K
; APPLICANT: Jiang, Xi
; APPLICANT: Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Ave., N.W.
; CITY: Washington, D.C.
; STATE:
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,049
; FILING DATE: June 7, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311.023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-486-049-3

Query Match 99.1%; Score 2793; DB 4; Length 530;
Best Local Similarity 98.7%; Pred. No. 1.9e-253;
Matches 523; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGVNPIDPWIIINFP 60
DB 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGVNPIDPWIIINFP 60

```

QY 61 VOAPQGETTISNNTPGGVLDLSLGLPHLNPFLHLSQMYNGWGNVRIMLAGNAFTA 120
DB 61 VOAPQGETTISNNTPGGVLDLSLGLPHLNPFLHLSQMYNGWGNVRIMLAGNAFTA 120
QY 121 GKIIIVSCIPPGGSHNLTTAQTLPFHVIADVRTLDPIEVPLEDRVNLFNHNRNQOTM 180
DB 121 GKIIIVSCIPPGGSHNLTTAQTLPFHVIADVRTLDPIEVPLEDRVNLFNHNRNQOTM 180
QY 181 RLVCMLYTPLRGTGGTGSFVAVGRVMTCPSPDFNPLFLVPTVEQKTRPFTLNLPLSS 240
DB 181 RLVCMLYTPLRGTGGTGSFVAVGRVMTCPSPDFNPLFLVPTVEQKTRPFTLNLPLSS 240
QY 241 LNSRAPLPISMGISPDNVQSVQFONGRCTLDGLRGVGTTPVSLSHVAKIRGTSNGTVIN 300
DB 241 LNSRAPLPISMGISPDNVQSVQFONGRCTLDGLRGVGTTPVSLSHVAKIRGTSNGTVIN 300
QY 301 LTELDTGTPHPPFGPAPIGFDPDLGGCDWHINMTQFGHSSQTYDVTTPDTFVPHLGSIQ 360
DB 301 LTELDTGTPHPPFGPAPIGFDPDLGGCDWHINMTQFGHSSQTYDVTTPDTFVPHLGSIQ 360
QY 361 ANGISGNYIGVLSWSPSPHSGSOVDLWKIPNYGSSITEATHLAPSVYPPGFGVLFV 420
DB 361 ANGISGNYIGVLSWSPSPHSGSOVDLWKIPNYGSSITEATHLAPSVYPPGFGVLFV 420
QY 421 FMSKIPGPGAYSLPCLLPQEIYISHLASEQAFTVGEAALLHYVDPDTRTLGFBKAYPDGF 480
DB 421 FMSKIPGPGAYSLPCLLPQEIYISHLASEQAFTVGEAALLHYVDPDTRTLGFBKAYPDGF 480
QY 481 LTCVPNGASSGPQOLPINGVVFVSVWSRYQLKPVGTASSARGRLGLR 530
DB 481 LTCVPNGASSGPQOLPINGVVFVSVWSRYQLKPVGTASSARGRLGLR 530

```

```

RESULT 2
US-09-617-594A-4
; Sequence 4, Application US/09617594A
; Patent No. 6541458
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALCIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.1
; CURRENT APPLICATION NUMBER: US/09/617,594A
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-617-594A-4

```

```

Query Match 8.7%; Score 244.5; DB 4; Length 668;
Best Local Similarity 32.2%; Pred. No. 4.3e-14;
Matches 69; Conservative 27; Mismatches 87; Indels 31; Gaps 8;
QY 19 GOLVPEVNASDPLAMPVAGSTAVATAGQVNPIDPWIINNFPVQAPQGEFTI-----SP 72
DB 136 GTLVGCVIAEPNAQMSAVD---VATGKSVDP--SEW-----EAFHFHSVNWST 180
QY 73 NNTPGGVLDLSLGLPHLNPFLHLSQMYNGWGNVRIMLAGNAFTAAGKIIVSCIPP 132
DB 181 SETQKILFKQSLGFLNPLNYLTHLAKLYVWSSIEVRFSISGSGVFGKLAIVVPPGI 240
QY 133 GSHNLTTAQTLPFHVIADVRTLDPIEVPLEDRVNLFNHNRNQOTMELVCMVLY 188
DB 241 DPVQST--SMLQPHVLFDAQVPEVIFTPDLRNSLYHLS--DITDTSLVIMINDLIN 297

```

```

QY 189 PLRTGGTGDSTFWAVGRVMTCPSPDFNPLFLVPP 222
DB 298 PYANDSNSGCIYV---VETKPGDPFKFHLKAPP 328
RESULT 3
US-09-590-020-2
; Sequence 2, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-2

```

```

Query Match 8.5%; Score 239; DB 4; Length 623;
Best Local Similarity 26.3%; Pred. No. 1.3e-13;
Matches 74; Conservative 43; Mismatches 98; Indels 66; Gaps 9;
QY 71 SPNTPGGVLDLSLGLPHLNPFLHLSQMYNGWGNVRIMLAGNAFTAAGKIIVSCIPP 130
DB 134 STSETQKILFKQSLGFLNPLNYLTHLAKLYVWSSIEVRFSISGSGVFGKLAIVVPP 193
QY 131 GFGSHNLTTAQT---LFPHVIAVRLDPIEVPLEDRVNLFNHNRNQOTMELVCMVLY 187
DB 194 G-----IEPVQSTMLQYPHVLFDAQVPEVIFAPDLRNSLYHLS--DITDTSLVIMVY 247
QY 188 TPLRT---GGTGDSTFWAVGRVMTCPSPDFNPLFLVPP----- 222
DB 248 NDILNPYANDTNSGCIY-TVETKPGDPFKFHLKAPPQSGMLTHGSPVSDLLIPKSSSLWIG 306
QY 223 -----TVEQKTRPFTLPNPLSLSSNRAPLPISGNGISPDN 259
DB 307 NRHMSDITDFTIIRPFVQANRHFDFNQETAGSTPRFRPITITVSESNMKLGIGVATDY 366
QY 260 -VQSVQFONGRCTLDGLRGVGTTPVSLSHVAKIRGTSNGTVI 299
DB 367 IVEGIPDGPDPDTTPEQL---TPAGIYSIT---ASNGTVI 400

```

```

RESULT 4
US-09-590-020-7
; Sequence 7, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 626

```


Db 248 NDLINPYANDTNSGCIIV-TVETKEGDPDKFHLKPP 283

Db 630 PLAPSDTVQTRHVQRRSRSESTIESFFARGACVAIIIVDN-----EQPTTRAKLFAM 683

Qy 318 --IGPDLGGCWHINMTQFGHSSOTQYDQVDTPTDTPFPHLGSIQANGISGNYIGVLSW 375

Db 684 WRITYKDT--VQIRKLEFFYS---RFDMEF---TFVVTANFTNANGHALNQVQIMY 735

Qy 376 VSPSPHSGQVLDLWKIPNYGSSITEATHLAPSVY 410

Db 736 I-PPGAPT-----PKSWDDYTQWTSNPSIF 760

RESULT 9

US-08-461-503-2

; Sequence 2, Application US/08461503

; Patent No. 5834302

; GENERAL INFORMATION:

; APPLICANT: Racaniello, Vincent

; APPLICANT: Tatem, Joanne M.

; APPLICANT: Weeks-Levy, Carolyn L.

; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES

; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10112

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/461,503

; FILING DATE: 5-JUN-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 36607-D-PCT-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 278-0400

; TELEFAX: (212) 391-0525

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2206 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-461-503-2

Query Match 4.6%; Score 129; DB 2; Length 2206;

Best Local Similarity 21.1%; Pred. No. 0.02;

Matches 96; Conservative 59; Mismatches 196; Indels 104; Gaps 20;

Qy 17 GAGQLVPEVNASDPLAMPVAGSSTAVATAGQVN-----PIDPWIIINNFVQAPQG---- 66

Db 349 GSNQYLTSNDHQSPCAI-PEFDVTPPIDIPGEVKNMELAEIDTMIPLNLESTKNTMDM 407

Qy 67 -EFTISPNNTGGVLFDSLGLPHLNPFLH-----LSQWYNGWGNMVRIMLAGNAFTA 120

Db 408 YRVTLSDSADLSQPLICLSLSPAFDPRLSHTMLGEVLNYTHWAGSLKFTFLFCGSMWAT 467

Qy 121 GKIIIVSCIPPGFGSHNLTIQAATLFPFHVIADVRTLDPIEVLPLEVDRVNLFHNNDRNQOT- 179

Db 468 GKILVAYAPPG-AQPPTSKEAMLGTHVWDLGLQSSCTMVVPVWISNVTYRQTQDSFTE 526

Qy 180 -----MRLVCMLYTPRLTGGTGDSSFVAGRVMTCPSPDENFLFL-----V 220

Db 527 GGYISMFYQTRIVVPLSTP-----KSMMLGFSVAC--NDFSRLLRDTHHSQSAL 576

Db 248 NDLINPYANDTNSGCIIV-TVETKEGDPDKFHLKPP 283

Db 630 PLAPSDTVQTRHVQRRSRSESTIESFFARGACVAIIIVDN-----EQPTTRAKLFAM 683

Qy 318 --IGPDLGGCWHINMTQFGHSSOTQYDQVDTPTDTPFPHLGSIQANGISGNYIGVLSW 375

Db 684 WRITYKDT--VQIRKLEFFYS---RFDMEF---TFVVTANFTNANGHALNQVQIMY 735

Qy 376 VSPSPHSGQVLDLWKIPNYGSSITEATHLAPSVY 410

Db 736 I-PPGAPT-----PKSWDDYTQWTSNPSIF 760

RESULT 8

US-07-852-260-2

; Sequence 2, Application US/07852260

; Patent No. 5525715

; GENERAL INFORMATION:

; APPLICANT: Racaniello, Vincent

; APPLICANT: Tatem, Joanne M.

; APPLICANT: Weeks-Levy, Carolyn L.

; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM

; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham

; STREET: 30 Rockefeller Plaza

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10112

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/852,260

; FILING DATE: 19920619

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 36607-B-PCT-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 977-9550

; TELEFAX: (212) 664-0525

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2206 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-07-852-260-2

Query Match 4.6%; Score 129; DB 1; Length 2206;

Best Local Similarity 21.1%; Pred. No. 0.02;

Matches 96; Conservative 59; Mismatches 196; Indels 104; Gaps 20;

Qy 17 GAGQLVPEVNASDPLAMPVAGSSTAVATAGQVN-----PIDPWIIINNFVQAPQG---- 66

Db 349 GSNQYLTSNDHQSPCAI-PEFDVTPPIDIPGEVKNMELAEIDTMIPLNLESTKNTMDM 407

Qy 67 -EFTISPNNTGGVLFDSLGLPHLNPFLH-----LSQWYNGWGNMVRIMLAGNAFTA 120

Db 408 YRVTLSDSADLSQPLICLSLSPAFDPRLSHTMLGEVLNYTHWAGSLKFTFLFCGSMWAT 467

Qy 121 GKIIIVSCIPPGFGSHNLTIQAATLFPFHVIADVRTLDPIEVLPLEVDRVNLFHNNDRNQOT- 179

Db 468 GKILVAYAPPG-AQPPTSKEAMLGTHVWDLGLQSSCTMVVPVWISNVTYRQTQDSFTE 526

Qy 180 -----MRLVCMLYTPRLTGGTGDSSFVAGRVMTCPSPDENFLFL-----V 220

Db 527 GGYISMFYQTRIVVPLSTP-----KSMMLGFSVAC--NDFSRLLRDTHHSQSAL 576

Qy 221 PPTVQKTRPFTPLNPLS-----SLNSRAPLPISGMISPDNVQSVQFQNGRCLTDG 274

Db 577 PGIGEDLTSVAQAGALTSLPKQDQSLPDTKA-----SGPAHSKEVPALTAVETG---ATN 629

Qy 275 RLVGTTPTVSLSHVAKIRGTSNGTV-----INLTDLGTGTPHPFEGPAP----- 317

QY 221 PPTVEQKTRPFTLPNPLS-----SLNSRAPLPISGMISPDNVQSVQFONGRCTLG 274
DB 577 POGIEDLTSEVAQAGALTSLPKQDLSLPTKA-----SGPAHSKEVPALTAETG---ATN 629
QY 275 RLVGTPPVSLSHVAKIRGTSNGTV-----INLTDLGTPFPFPFEGPAP----- 317
DB 630 PLAPSDTVQTRHVQRRSRSESTISFFARGACVAIIEDVN-----EQPTTTRAKLFPAM 683
QY 318 --IGPPDLGGCDWHINMTQFGHSSQTDVDTPTTFVPHLGSIOANGIGSGNYIGVLSW 375
DB 684 WRITYKDT--VQLRRKLEFFTYS---REFDMEF---TFVVTANFTNANNCHALNQVQIWMY 735
QY 376 VSPSPHPSGSQVDLWKIPNYGSSITEATHLAPSVY 410
DB 736 I-PPGAPT-----PKSWDDTYWTQSSNPSIF 760

RESULT 10

US-08-465-250-2
; Sequence 2, Application US/08465250
; Patent No. 6136570
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release 1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,250
; FILING DATE: 6-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-B-PCT-US
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-465-250-2

Query Match 4.6%; Score 129; DB 3; Length 2206;
Best Local Similarity 21.1%; Pred. No. 0.02;
Matches 96; Conservative 59; Mismatches 196; Indels 104; Gaps 20;

QY 17 GAGQLVPEVNASDPLAMPDVAGSSTAVATAGVN-----PIDPWIINNFVQAPG----- 66
DB 349 GSNQVLTSDNHQSPCAI-PEFDVTPPIDPGEVQNMELAEIDTWIPLNLESTKENTMDM 407
QY 67 -EFTISPNNTPGGVFLDLSLGLHNPFLH-----LSQWYNGWGNMVRIMLAGNAFTA 120
DB 408 YRVLTSDSADLSQPLICLSLSPAFDRLSHTMLGEVLNYYTHWAGSLKFTFLFCGMMAT 467

QY 121 GKLIYSCIPPGSGSHNLTIQAATLPPHVIADVRTLDPIEVPLEDVRNVLPHNDRNQQT- 179
DB 468 GKILVAYAPPG-AQPTSRKEAMLGTHVWIDLGLOSSCTMVVPMWISNTYRTQTDSPTE 526
QY 180 -----MRLVCMLYTPLRTGGGTGDSFVVAGRVMTCPSPDFNLFEL-----V 220
DB 527 GGYISMPFQTRIVVPLSTP-----KSMMLGFSVAC--NDPSVRLLRDTHHSQSAL 576
QY 221 PPTVEQKTRPFTLPNPLS-----SLNSRAPLPISGMISPDNVQSVQFONGRCTLG 274
DB 577 POGIEDLTSEVAQAGALTSLPKQDLSLPTKA-----SGPAHSKEVPALTAETG---ATN 629
QY 275 RLVGTPPVSLSHVAKIRGTSNGTV-----INLTDLGTPFPFPFEGPAP----- 317
DB 630 PLAPSDTVQTRHVQRRSRSESTISFFARGACVAIIEDVN-----EQPTTTRAKLFPAM 683
QY 318 --IGPPDLGGCDWHINMTQFGHSSQTDVDTPTTFVPHLGSIOANGIGSGNYIGVLSW 375
DB 684 WRITYKDT--VQLRRKLEFFTYS---REFDMEF---TFVVTANFTNANNCHALNQVQIWMY 735
QY 376 VSPSPHPSGSQVDLWKIPNYGSSITEATHLAPSVY 410
DB 736 I-PPGAPT-----PKSWDDTYWTQSSNPSIF 760

RESULT 11

US-08-689-421-33
; Sequence 33, Application US/08689421
; Patent No. 6008029
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie S.
; APPLICANT: Brown, Kimberley M.
; APPLICANT: Kauppinen, Sakari
; APPLICANT: Halkier, Torben P.
; TITLE OF INVENTION: PURIFIED COPRINUS LACCASES AND NUCLEIC ACIDS
; TITLE OF INVENTION: ENCODING SAME
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 60080290 No. 6008029disk of No. 6008029th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/689,421
; FILING DATE: 9-AUG-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Harrington, James J.
; REGISTRATION NUMBER: 38,711
; REFERENCE/DOCKET NUMBER: 4554.204-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 517 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-689-421-33

Query Match 4.5%; Score 127; DB 3; Length 517;
Best Local Similarity 23.8%; Pred. No. 0.003;
Matches 124; Conservative 59; Mismatches 200; Indels 138; Gaps 30;

QY 40 STAVATA-----GOVNPIDPWIINNFVQAPQGEFTIS-----PN-NTPGGVL----- 80

TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 517 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-389-528-33

Query Match 4.5%; Score 127; DB 3; Length 517;
Best Local Similarity 23.8%; Pred. No. 0.003;
Matches 124; Conservative 59; Mismatches 200; Indels 138; Gaps 30;

5 ATALATSLPFLVGLGAGSTNLVANKVAPDG-FRSRAVLGATQTPVQPGVQGNK 63
81 ---FDLSLGPLH-NPFLHLQMYNGWGNVRIMLAGNAFTAGKIIVSCIPPGFGSHN 136
64 NSFFAINVIDALTDPTMLRTTSH--WHG-----MFQGTAWADGAGVTCQIPSPGHSP 116
137 LTIAQA-----TLFPHVIADVRTLDPIE-----VPLEDVRNVLFHNDRNQOTMRLVCM 185
117 LYKFOALNOAGTFTWYSHHESQYCDLRCAMVVYDVPDHRNL--YDIDNEATITLADW 174
186 LYTPRLRTGG--GTGDSFVAVAGVMTCPSPDFNLFVLPPTVEQKTR-----PFTL----- 233
175 YHVPAAPSAGLVPTPDSTLINGKRYAGGP-----TVPLAVISVTRNRRYRFRVLSLSCD 228
234 PNLPLSSLSNRAPLPISGMGI SPDNVQSVQF-----QNGRGT 271
229 PNYVFSIDGHTMTVIEVDGNNVQPLVDSIQIFAGORYSFVLNANRPVGNVYVRANPNIG 288
272 LDGRLVGTTPVSLSHVAKIRGTSNGTVINLTLDGTFP-----HPFEGPAPIGPPD 322
289 TTGFVGVNSAILRYV---GASN-TDPTTQ---TPFSNPLETLNHLPLTNPAAPGLPT 340
323 LGGCDWHINM-TQGHSSQTYDVTPTDF-----VPHLGSIQANG-----IGSGN-YI 370
341 PGGVDVAINTVDFSSLT-FSVNGA--TFHQPPVPVLLQIMSGAQTAAQLLPSSGVYV 397
371 GVLSWVSPSPHSGSQVDLWKIPNYGSSITEATHLAPSVYPPGFG-----QNGRGT 415
398 LPRNKVIELSNPGGSTGSPHFLHGHEFAVVRSGASTY--NFANPVRDVRVSVAGVD 455
416 EVLVFFMSKIPGPGAYSLPCLLPQEIYSHLA---SEQAPTV 453
456 NVTIREFRTDNFGP--WILHCHIDWHLVGLAVVFAEDAPT 494

RESULT 13
US-09-389-528-33
Sequence 33, Application US/09389528
Patent No. 6207430
GENERAL INFORMATION:
APPLICANT: Yaver, Debbie S.
APPLICANT: Brown, Kimberley M.
APPLICANT: Kauppinen, Sakari
APPLICANT: Halkier, Torben P.
TITLE OF INVENTION: PURIFIED COPRINUS LACCASES AND NUCLEIC ACIDS
TITLE OF INVENTION: ENCODING SAME
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESS: No. 62074300 No. 6207430disk of No. 6207430th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/389,528
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/689,421
FILING DATE: 9-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Harrington, James J.
REGISTRATION NUMBER: 38,711
REFERENCE/DOCKET NUMBER: 4554.204-WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 517 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-389-528-33

Query Match 4.5%; Score 127; DB 3; Length 517;
Best Local Similarity 23.8%; Pred. No. 0.003;
Matches 124; Conservative 59; Mismatches 200; Indels 138; Gaps 30;

40 STAVATA-----GOVNPDIPINNFFVQAPQGEFTIS-----PN-NTPGVVL----- 80
5 ATALATSLPFLVGLGAGSTNLVANKVAPDG-FRSRAVLGATQTPVQPGVQGNK 63
81 ---FDLSLGPLH-NPFLHLQMYNGWGNVRIMLAGNAFTAGKIIVSCIPPGFGSHN 136
64 NSFFAINVIDALTDPTMLRTTSH--WHG-----MFQGTAWADGAGVTCQIPSPGHSP 116
137 LTIAQA-----TLFPHVIADVRTLDPIE-----VPLEDVRNVLFHNDRNQOTMRLVCM 185
117 LYKFOALNOAGTFTWYSHHESQYCDLRCAMVVYDVPDHRNL--YDIDNEATITLADW 174
186 LYTPRLRTGG--GTGDSFVAVAGVMTCPSPDFNLFVLPPTVEQKTR-----PFTL----- 233
175 YHVPAAPSAGLVPTPDSTLINGKRYAGGP-----TVPLAVISVTRNRRYRFRVLSLSCD 228
234 PNLPLSSLSNRAPLPISGMGI SPDNVQSVQF-----QNGRGT 271
229 PNYVFSIDGHTMTVIEVDGNNVQPLVDSIQIFAGORYSFVLNANRPVGNVYVRANPNIG 288
272 LDGRLVGTTPVSLSHVAKIRGTSNGTVINLTLDGTFP-----HPFEGPAPIGPPD 322
289 TTGFVGVNSAILRYV---GASN-TDPTTQ---TPFSNPLETLNHLPLTNPAAPGLPT 340
323 LGGCDWHINM-TQGHSSQTYDVTPTDF-----VPHLGSIQANG-----IGSGN-YI 370
341 PGGVDVAINTVDFSSLT-FSVNGA--TFHQPPVPVLLQIMSGAQTAAQLLPSSGVYV 397
371 GVLSWVSPSPHSGSQVDLWKIPNYGSSITEATHLAPSVYPPGFG-----QNGRGT 415
398 LPRNKVIELSNPGGSTGSPHFLHGHEFAVVRSGASTY--NFANPVRDVRVSVAGVD 455
416 EVLVFFMSKIPGPGAYSLPCLLPQEIYSHLA---SEQAPTV 453
456 NVTIREFRTDNFGP--WILHCHIDWHLVGLAVVFAEDAPT 494

RESULT 13
US-09-181-827A-33
Sequence 33, Application US/09181827A
Patent No. 6242232
GENERAL INFORMATION:
APPLICANT: Yaver, Debbie S.
APPLICANT: Brown, Kimberley M.
APPLICANT: Kauppinen, Sakari
APPLICANT: Halkier, Torben P.
TITLE OF INVENTION: Purified Coprinus Laccases And Nucleic
TITLE OF INVENTION: Acids Encoding Same
FILE REFERENCE: 4554.200-US
CURRENT APPLICATION NUMBER: US/09/181,827A
CURRENT FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/002,800
PRIOR FILING DATE: 1995-08-25
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 33
LENGTH: 517
TYPE: PRT
ORGANISM: Coprinus cinereus
US-09-181-827A-33

Query Match 4.5%; Score 127; DB 3; Length 517;
Best Local Similarity 23.8%; Pred. No. 0.003;
Matches 124; Conservative 59; Mismatches 200; Indels 138; Gaps 30;

QY 40 STAVATA-----GQVNPIDPMLNNFVQAPQGEFTIS-----PN-NTGGVL-----80
DB 5 ATALATSLLPFLVGLGAPSTNLVANKVIAPDG-FSRSAVLAGATQPTVQFPVPGVIOGNK 63
QY 81 ----FOLSLGPHL-NPFLHLSQMYNGWGNMVRIMLAGNAFTAGKIIIVSCIPPGFGSHN 136
DB 64 NSFFAINVIDALTDPTMLATTSIH--WHG-----MFQGTAWADGAGVTCQIPSGHSF 116
QY 137 LIAQA-----TLFPHVIAVRLTDPIE-----VPLEDRVNLVFNHNDNNOQTWRLVCM 185
DB 117 LYKFOALNOAGTFWYHSHHESQYCDGLRGAMVYDVPDPRNL--YDIDNEATIITLADW 174
QY 186 LYTLRTGG--GTGDSFVAGRVMTCPSPDNFLFLVPTVEQKTR-----PPTL-----233
DB 175 YHVPASAGLVPTPDSLTNGRYAGGP-----TVPLAVISVTRNRRYRFLVLSLSD 228
QY 234 PNLPLSSLSNSRAPISPGMISGSPONVQSVQF-----QNGRCT 271
DB 229 PNVFSIDGHTVTVIEVDGVNQPLVDSIQIFAGORYSFVLNANRPVGNVVRANPNIG 288
QY 272 LGRLVGTTPVSLHSHVAKIRGTSNGTVINLTLDGTPF-----HPFEGPAPIGPPD 322
DB 289 TTGFVGGVNSAILRYV---GASN-TDPTTQ---TPFSNPLETLNHLPLTNPAAPGLPT 340
QY 323 LGGDWHINM-TQFQSSQTDVDTPTDF-----VPHLSIQANG-----IGSN-YI 370
DB 341 PGGVDVAINLNTVFPSSLT-FSVNGA--TFHQPPVPLLQIMSGAQTAAQLLPSGSVYV 397
QY 371 GVLWSVPPSPHSPSGQVLDLWKPINYGSSITEATHLAPSVYPPGFG-----415
DB 398 LPRNKVIELSMGGSTGSPHPLHGHFPAVRSAGSTY--NFANPVRDVVSAGVAD 455
QY 416 EVLVFMSKIPGAYSLPCLLPQEIYISHLA---SEQAFTV 453
DB 456 NVTIRFRDTPNGF--WILCHIDHVLVLGLAVVFAEDAPT 494

RESULT 14

US-09-328-352-5503
; Sequence 5503, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5503
; LENGTH: 3892
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5503

Query Match 4.4%; Score 124.5; DB 4; Length 3892;
Best Local Similarity 21.0%; Pred. No. 0.13;
Matches 125; Conservative 70; Mismatches 243; Indels 157; Gaps 31;

QY 7 DATSSVDGAGAGQLV-PRVNASD---PLAMPDVAGSS-TAVATAGVNP-----ID 53
DB 622 DTATVAGPDGWSVPNGLNDGVEAATDPAGNPSLPGTATVDAGPNTDGVNFTVD 681
QY 54 PWIINNFOAPQGEFTISPNPTPGVLFDSLGLPHLNPFLHLSQMYNGWGNMVRIML 113
DB 682 SVTADNVINASEA-----SGNVTVTGVLKNVPAD-----AANTVVTVVI 720
QY 114 AGNAFTAGKIIIVSCIPPGFGSHNLTAQATLFPFHVADV-RTLDPFIEVPLEDRVNLVFNH 172

DB 721 NGQTYTA-----TVDSSTAGTWTVSPGSDL-----TADADKTID-AKVTFTDAAGNSSSV 769
QY 173 NDNRNQOTMRLV---CMLVTPLR-----GGT-----GDSFVWAG--RVMTCP 210
DB 770 NDTQTYTIDTAPDAPVINPVNGTDPITGTAEPGSTVTVYPDGSTTVVAGPDGTWTV 829
QY 211 SPDFNFLVPTVBQKTRPFTLPNLPLSSLSNSRAP-----LPISGMGISPNVQSVQFO 266
DB 830 NPLGNDGDKVTAIA---TDPAGNPSLPGTATVDAGPNTDGVNFTVDSVTADNVINASEA 886
QY 267 NGRCTLDGRL-----VGTTPVSL-----SHVAKIRGTSNGTVINLTDELDTGTFPFPFEGP 315
DB 887 SGNVTVTGVLKNVPADAANTVTVVINGQTYTATVDSSTAGTWTVSPGSDLTADADKTID 946
QY 316 APIGPDPDGGCDWHINMTQFGHSSQTDVDTPTDFVPHLGSIQANGIGSGNYIGVLSW 375
DB 947 AKVTFTDAAGNSSSVNDTH-----TYTVDVAPN-APVLDPINATDPVSGQ-----991
QY 376 VSPPS-----HPSGSQVLD-----WKIPNYGSSITEATHLAPSVYPPGGEVLVFFMS 423
DB 992 AEPGSTVTVYPDGTTATVAGPDGSGWSVPNPNGLVDGTVTATATDPAG-----NT 1043
QY 424 KIPGCGAYSLPCLLPQEIYISH-LASEQAPTVGEAALLHYVDDPTGRTLGEFKA--YP---477
DB 1044 SLPGTGTVSADITAPVVALDDVLTNDSTP-----ALTGTVNDPTATVVNVVDGVDPAYN 1098
QY 478 --DGLTCTCPN---CASSGPQQLPINGVVFVWSVRSFQVKPKPVATSSARGLG 527
DB 1099 NGDGTWTLADNLTPLTDPGPHITIV-----TATDAAGNAG 1133

RESULT 15

US-08-746-111-5
; Sequence 5, Application US/08746111
; Patent No. 6066778
; GENERAL INFORMATION:
; APPLICANT: Ginsburg, David
; APPLICANT: Cui, Jisong
; TITLE OF INVENTION: Compositions And Methods For Screening
; TITLE OF INVENTION: Compounds For Anticoagulant Activity
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/746,111
; FILING DATE: 06-NOV-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: UM-02536
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2183 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-746-111-5

Query Match	4.4%	Score 123.5;	DB 3;	Length 2183;
Best Local Similarity	26.8%	Pred: No. 0.063;		
Matches 82;	Conservative 31;	Mismatches 110;	Indels 83;	Gaps 18;

QY	212	PDFNLFVPPTEQKTRPFTLPNLPSSLSNSRAPLPISG-MGISPDNVQSVQFQNGRC	270
Db	1196	FLDDQL-LLSPEDNQKTSPPDLGQVPLSPDDNQKTSPPD---	1251
QY	271	TLDGRIVGTFVSLSHVAKIRGTSNGTVINLTDLGTFPHPEGPAPIGFPDGLGCCDWHI	330
Db	1252	-----LGQVPLSDDDNQKTSPPDLGQV-----PLSPDDNQK-ITSPDLG-----	1289
QY	331	NMTQFGHSSQTYDVTTPDTFVPHLGSIQANGISGNYIGVLSWV--SPESHPSGQVD	388
Db	1290	----QVPLSSDNQ-----KTSSPDLGQVPLFPEDNQNYFLDLSQVPLSSDNQKTSSTD	1339
QY	389	LWKI-PNYG-----SSITEATHLAPSVYPPGFEVLVFFMSKIPGGA	430
Db	1340	LLTSLSPDFGQTVLSPDLQDLPLPSDNSQVTVSPDLSLTLSPDFNEIIL-----APDLGQ	1394
QY	431	YSL-PCLL---PQYIYSHLAS-----EOA---PTVGEAALLHYVDPTGRTLGFEKAYPDG	479
Db	1395	VTLSPDLIQTNPALNHGHKASSADPDQASYPDPDSGQASSL-----PELNRTL-----PHPD-	1445
QY	480	FLTCVP	485
Db	1446	-LTHIP	1450

Search completed: June 1, 2004, 13:57:54
Job time : 14.6275 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 33.8809 Seconds
(without alignments)
4368.312 Million cell updates/sec

Title: US-09-926-799-2
Perfect score: 2819
Sequence: 1 MMASKDTSSVDGASGAGQ.....YOLKPVGTASSARGRLGRR 530

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2793	99.1	530	14	US-10-314-739-3
2	244.5	8.7	668	14	Sequence 3, Appli
3	238	8.4	669	14	Sequence 4, Appli
4	237	8.4	547	12	Sequence 2, Appli
5	237	8.4	671	12	Sequence 2, Appli
6	126.5	4.5	1049	12	Sequence 49900, A
7	126	4.5	6310	12	Sequence 67793, A
8	124.5	4.4	913	15	Sequence 2294, Ap
9	121.5	4.3	749	12	Sequence 55747, A
10	118.5	4.2	8360	14	Sequence 34, Appl
11	116	4.1	2468	14	Sequence 66335, A
12	116	4.1	2468	14	Sequence 4, Appli
13	112.5	4.0	878	14	Sequence 171, App
14	112	4.0	702	12	Sequence 366, App
15	112	4.0	769	12	Sequence 212925, A

16	112	4.0	925	15	US-10-108-260A-3752	Sequence 3752, Ap
17	111.5	4.0	1765	12	US-10-282-122A-48055	Sequence 48055, A
18	111.5	4.0	6620	15	US-10-080-334-290	Sequence 290, App
19	111	3.9	1439	12	US-10-282-122A-44567	Sequence 44567, A
20	108.5	3.8	651	15	US-10-346-144-2	Sequence 2, Appli
21	108.5	3.8	769	14	US-10-029-386-3366	Sequence 3366, A
22	108	3.8	684	12	US-10-423-114-43059	Sequence 43059, A
23	108	3.8	724	15	US-10-346-144-5	Sequence 5, Appli
24	108	3.8	1465	15	US-10-369-493-13955	Sequence 13955, A
25	108	3.8	4952	15	US-10-051-874-56	Sequence 56, Appl
26	108	3.8	5159	15	US-10-085-198-112	Sequence 112, App
27	108	3.8	5262	15	US-10-051-874-165	Sequence 165, App
28	107.5	3.8	1048	14	US-10-223-538-10	Sequence 10, Appl
29	107	3.8	2358	12	US-10-282-122A-45763	Sequence 45763, A
30	106.5	3.8	668	12	US-10-282-122A-72916	Sequence 72916, A
31	106	3.8	792	15	US-10-369-493-4878	Sequence 4878, Ap
32	106	3.8	792	15	US-10-369-493-7637	Sequence 7637, Ap
33	106	3.8	3930	12	US-10-282-122A-46817	Sequence 46817, A
34	105.5	3.7	750	9	US-09-815-242-13985	Sequence 13985, A
35	105.5	3.7	750	12	US-10-282-122A-75580	Sequence 75580, A
36	105.5	3.7	752	12	US-10-282-122A-59788	Sequence 59788, A
37	105.5	3.7	996	14	US-10-101-464A-889	Sequence 889, App
38	105.5	3.7	996	14	US-10-101-464A-933	Sequence 933, App
39	105	3.7	407	14	US-10-156-761-12940	Sequence 12940, A
40	105	3.7	599	10	US-09-955-555A-29	Sequence 29, Appl
41	105	3.7	5262	15	US-10-051-874-167	Sequence 167, App
42	104.5	3.7	785	15	US-10-369-493-11883	Sequence 11883, A
43	104.5	3.7	5008	15	US-10-051-874-166	Sequence 166, App
44	103.5	3.7	353	10	US-09-946-374-397	Sequence 397, App
45	103.5	3.7	353	12	US-10-006-485A-397	Sequence 397, App

ALIGNMENTS

RESULT 1
US-10-314-739-3
; Sequence 3, Application US/10314739
; Publication No. US20030129588A1
; GENERAL INFORMATION:
; APPLICANT: ~~Genentech, Inc.~~
; Jiang, Xi
; Graham, David Y

TITLE OF INVENTION: Methods and Reagents to Detect and Characterize No. US20030129588A1walk and Related Viruses

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski L.L.P.
STREET: 801 Pennsylvania Ave., N.W.
CITY: Washington, D.C.
STATE: <Unknown>
COUNTRY: USA
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/314,739
FILING DATE: 09-Dec-2002
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/486,049
FILING DATE: June 7, 1995

ATTORNEY/AGENT INFORMATION:
NAME: Davis, Peter
REGISTRATION NUMBER: 36,119
REFERENCE/DOCKET NUMBER: 311.023
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-662-0200
TELEFAX: 202-662-4643
TELEX: <Unknown>

43 -----VATAGVNPIDPWIINNFAQAP-----QG 66
383 VCANANIITAGSTIITDNNGVLLSTGSPSTTTDFDAGASTLGAITTGAGTTTTLHDDT 442
67 EFTISPN-----NTPGGVLEDLGLHLPFLHLS-OMYNGWGVNMVRIMLAGNAF 118
443 DMTVGNSTATNLTSASNIHTAPADPSQAASYKTLOVQNTGADGNLHNLTLGSDDA 502
119 TAGKIVSCIPPGFSGHNLITIAQATLFPHVIAADVRLDPIEVLPELDBVRNLFHNNDRNQ 178
503 PSDKLIVSG-GRATGTNLTIHNTT-----GPGDMIVADGIQV-----VQ 541
179 TMLRVCLMPLYLRTGGT--GDSFVAVAGVMTCPSPDNFL----- 217
542 TL-----NGGTTASDAFTLTGLR---AGAFDYLLFHGGIGGSRPNDFLRSS 586
218 FLVPTVEOKTRPFTLNLPLSLNSRAP---LPISGMGISPNVQSVQFQNGRCRTLD 273
587 FVVPESI-----TPTVPLSPATSPSPGEGTSPSPG-EPGTPSPPGQCTPSSP 637
274 GRLVGT-----TPVSLSHVAKIRGT-----SNGTVINLTELD 305
638 PGQGTTPSPPGQPGTSPSPGQPGTSPSPGQPGAPSPSPGQPGTSPSPGQPGQTTLD 697
306 -GTPP-----HPPEGP-----APIGPDLGGCDWHINMT----- 333
698 VSRFPFADPPASPLOQGVTVFVIGELATYGVVQVVARQWFTMLGTLHERIGDILTADN 757
334 -----QFGHSSQTOY---DVTTPDTFVPHLGSIQANGIGSGNYIGVLSWSPSPSHPSG 384
758 AGPDSGSHGSHWARFPGQIDNRYQAFASPRATGQLLGVQAG----- 800
385 SQVDMK---IPN-----YGS-----ITEATHLAPSVYPGEGEVLPFFMSK 424
801 --VDVWRGSLMPGRDVRAGVYFAYGNSNLVDVGLVTNATATA-----YLLGRGK 848
425 IP-----GPGVSLPCLLPQEVYSHLASEQAPTVGEAALLHYVDVDPDTGRTLG 471
849 LNLDAYSGGAYWTHYGGWGLDLVLOG-----TSVQGEATVQSARL-----PTTG--- 894
472 EFKAYPBGFTLCPNGASSGQQPLNGVVFVWSWSRFYQ-----LKPV---GTA 519
895 -----DGFVTSLEAGY---PVPPLGLFRFVLEPQIQVLMQVRVFDQNDGLGFLVGLST 945
520 SSARGRLGLR 529
946 TGTGRLGLR 955

RESULT 7
US-10-282-122A-67793
; Sequence 67793, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; FILE OF INVENTION: Identification of Essential Genes in Microorganisms
; TITLE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848

;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/230,335
;; PRIOR FILING DATE: 2000-09-06
;; PRIOR APPLICATION NUMBER: 60/230,347
;; PRIOR FILING DATE: 2000-09-09
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/267,636
;; PRIOR FILING DATE: 2001-02-09
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 78614
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 67793
;; LENGTH: 6310
;; TYPE: PRT
;; ORGANISM: Pseudomonas putida
US-10-282-122A-67793

Query Match 4.5%; Score 126; DB 12; Length 6310;
Best Local Similarity 20.1%; Pred. No. 0.62;
Matches 122; Conservative 69; Mismatches 186; Indels 230; Gaps 31;

QY 6 KDATSSVDGAS---GAGQL-----VPE-----VNASDPLA-----MDPVAGSST--AV 43
Db 3278 RDADGNVLGTATVGGDGRFDTLNAPORNGESITVEATDNLGNSAGPVDFTAPDSITPPQV 3337
QY 44 ATAGOVNPIDPWIINN-----FVQAPQGEFTISPNTGEGVLFDSLGL-PHNLNPELLH 95
Db 3338 VTNPAIDIGITITNGEPEGATVTVRPGDGVLSGVVAGGA-FEITLDEPQTNGOALT 3396
QY 96 LSQMYNGWGVNMVRV-----MLAGNAFTAGKIIVSCIPG-----FGS 134
Db 3397 VEQ--RDAAGNLSAAVLSAPDTPQAPDPTGLVLADG-----GSSLSSGSGFAGVTVTGT 3450
QY 135 HNLITIAQATLPHVIADVRLDPIEVLPELDBVRNLFHNNDRNQTMRLVCLMYTLPRTGG 194
Db 3451 GGAVLGTATVQPDGTFQV-AIDPPQL-----NGQTLSTV-----QADNSGN 3490
QY 195 GTGDSFVVAVAGVMTCPSPDNFLVLPPTVEQKTRPFTLNLPLSSLSNSRAPLPISGMG 254
Db 3491 GSPAGTVIA-----PDF-----EAPLPAQGLG 3512
QY 255 ISPDNV-----QSQVFQNGRCTLGRLVGTTPVSLS---HVAKIRGTSNGTVINLT 302
Db 3513 LDPTGTLSSGQGEAGSTVEVRNAA-----GNLLGITPVNADGTFEVALSPAQTNGEVLVSV 3568
QY 303 ELDGTPFPFPEGAPIGFPDLGGCDWHINMTQFGHSSQTOYVDVDTTPD-----TFVPHLG 357
Db 3569 LIDG-----GG-----NASPSATFTADDSTAPAAPAGLTIPTGCS 3603
QY 358 SIOANG-----IGSGNYIGVL-----SWVSP--PSHPSSGQVDLWKIPNYGSSIT 400
Db 3604 AIOGSGEAGSTVEVKLADGTLVGTVVVAGGSFTVPLSPAQLDQALNVLTLDAAGNISQ 3663
QY 401 EATHLAPSVYPP-----GFEVLVFFVMSKIFPGPAY 431
Db 3664 PSQIFAPDITPPALPTDVAVSSDGTAVTGNAPGASSVTVSDGAGNVITVAVNP---DGSF 3720
QY 432 SLFCLLPQBYISHLASEQAPTV-----GEAALLHYVDVDPDTGRTLGEFKAYPDGFLTCV 484
Db 3721 SVFLDTPQN-----NGQTVTVVVTDAAGNDSAPSVTAPDT-----TNPEPATGLTVS 3768
QY 485 PNGASSG 491
Db 3769 PDGSTVG 3775

RESULT 8

US-10-369-493-2294
; Sequence 2294, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2294
; LENGTH: 913
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(913)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-2294

Query Match 4.4%; Score 124.5; DB 15; Length 913;
Best Local Similarity 19.6%; Pred. No. 0.044;
Matches 111; Conservative 65; Mismatches 197; Indels 193; Gaps 26;
QY 18 AGQVPEVNASPLAMDPAVAGSTAVAT-----AG--QVNPIDPWIINNFPQPGCF 68
Db 16 AGTVVNVQVEPVEVTVPAVGTITSTVPEVVPXAGIVEVNGIDPETIVSVTVPAVGT 75
QY 69 TISPNNTPGGVLDLSGLPHLPFLHLHLSQMYNGVGNMVRIMLAGNAFTAGKIIVSCI 128
Db 76 T-----STVPDVVPX-----AGNVEVVP-----TEPDIMVSVT 103
QY 129 PGFGSHNLTIAQTLFPHV-----IADVRLDP-----IEVPLEVRNVLPHNDRNQ 178
Db 104 VPAVGSTTSTV-----PDVXPAGIVEVNGIDPETIVSVTVPAVCA----- 144
QY 179 TWRLVCMVLTPLTGGTGGDSFVAGVMTCPSPDENFLVPPVTEQKTRFTLPNLP 238
Db 145 -----TISTVPDVVPXAGNVEVYPTETPDIIIVSVTVPAVGSTTS--TVPDV-- 187
QY 239 SSLNSRAPLPISGM-----GISPDNVQSVQFQNGRCTLDGLVGTTP----- 281
Db 188 -----VPXAGIVEVNGIDPETIVSV-----TVPAVATISTVDPVVLXAGNVEVVP 233
QY 282 -----VLSHVAKIRGTSNGTVINLTLDG--TFPHPEGPAPIGFDPDLGGCDWHNNMQ 336
Db 234 TEPDIMVSVTVPAVGSTTSTVDPVVPXAGNVEVYPT-----PDI-----MVSVTVA 281
QY 337 HSSQTYDYDTTDPDFVPHLSIQANGICSGNY-----IGVLSVSPSPSPSSQVDL 389
Db 282 VGSST-----STVPDV--VPXAGIVEVNGIDPETIVSVTVPAVGSTTMTAPDEVSAVINDI 336
QY 390 WK-----IPNYGSSITEATHLAP-----SVYP-----PGFGEVLVFFMS 423
Db 337 KETDPEITVSVTVPAVGSTTSTVTVPLAGNVEVYPTETPMVSVIVPAVGSTTSTVPE 396
QY 424 KIPGGAISLPLPOEYISHL-----ASEQAPTGVBAALHYVDPTGRTLGBF 473
Db 397 VVPLAGNVEVKVLEPETMVSIVPAVGSTTSTVPEVVPVPLAGNVE--VNALEPDITVSV-- 452
QY 474 KAYPDGFLTCVNGASSGQQPLPING 499
Db 453 -----IVPAVGSTTSTVPEVVPVLAG 472

RESULT 9

US-10-282-122A-55747
; Sequence 55747, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-03-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55747
; LENGTH: 749
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-55747

Query Match 4.3%; Score 121.5; DB 12; Length 749;
Best Local Similarity 21.3%; Pred. No. 0.062;
Matches 120; Conservative 54; Mismatches 174; Indels 215; Gaps 33;
QY 28 SDPLAMPV-AGSSTAVTAGQVNPIDPWIINNV---QAPQGETISPNTPGGVLPDL 83
Db 150 SDPKITPVVRFSTVQGGAGSADTVRD--IRGFATKFTYEGIFDLVGNNTPVFFIQD- 206
QY 84 SLGPHLNPFLHLHLSQMYNGVGNMVRIMLAGNAFTAGKIIVSCIPPGFGSHNLTIAQAT 143
Db 207 ---AHKFPDFVHAVKPEPHM-----AIPQGSADH-----T 234
QY 144 LPPHVIADVRLDPIEVEPLEDRNVLPHNDRNQOTMLRVLCTMLYPLTGGGTG----- 197
Db 235 FMDYV-----SLQP-----ETLHNVWVWMSDRG-----IPRSYRTMEGFGIHTPL 275
QY 198 -----DSFV-----VAGRV-----MTCPSPDFN-----PLF 218
Db 276 INAEKATFVFWKVPAGKASLVWDEAQKLTGRDPDFHRELWESIEAGDPPELVGLQ 335
QY 219 LVPPTVEQKTRPFTLPNL-PLSSLSNSRAPLPIG---MGISPDNV-----QSVQFQNGRC 270

314 ---GPAPIGPDIIGDGMHNMTOF---GHSSQTVQVDVDTPTDTPFVHLGSIQANG---I 364
1714 RVFGDAPAGSVVLGSKVTNLGHTQPAAGIAGVLKALLALQEQRVPSLHFAANARVPLD 1773
365 GSNYIGVLG---WVSPSPSHSGSVDLWIKIPNYGSSITEATHLAPSVYPP----- 412
1774 GSPFTVATTAAQWPEPAEGPRRAVS-----SFGASGTNA-HVLEEHPPVTRATTGPESA 1827
413 -GGEVLVFFMSKIPG-----PG-----AYSLPC----- 435
1828 GGDGEAAFLASPTAALRAVAERLLAREREGSLPARQVAYSLSAAGRRHFPHRLAVVAT 1887
436 ---LLPOEYISHLASEQAPTGVGEAALLHYVDPTGTR--TLGEF-----KAYPDG-----F 480
1888 GLPALAARLARLAWLADEQPG--GEGTLLHGVHAHAGTRQAALGGLAPAEALAAAYVGGAEPPF 1945
481 LTCVPNGCASSGPQQLPI 497
1946 AESFPAGAR---RQVPL 1959

RESULT 11
US-10-282-122A-66335
; Sequence 66335, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66335
; LENGTH: 2468
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-66335

Query Match 4.1%; Score 116; DB 12; Length 2468;
Best Local Similarity 22.2%; Pred. No. 1.3;

336 LIPEDEFEK---FDFDLDDPTKLIPEELVPQLVGMVLRNPNDFFAENQVAFHPGHI 392
271 T-----LDGLRGVTTVPVSLGVAKIRGTSNGTIVNLTELDTGTFPHFPEGPAITGF 320
393 VEGLOFTNDPLQGLRFSYDTQIS-----RLGGPNFHEI-----PINR 431
321 POLGGCDMHNMTOFCHSSQTVQVDVDTPTDTPFVHLGSIQANGSIYIGVLWV---VSP 378
432 PT---CPYH---NFQDGMHQDIDTNPANYEPN--SINDN-----WPRETP 470
379 PSHPSG-----SQVDLWIKIPNYGSSITEATHLAPSVYPPGGEVL-----VFEMSKIP-- 426
471 PGKRGGGSYQERID-----GNKIRERS-----PSGEYVAHPRFWSQTFIE 515
427 -----GPGAYSLPCLLPQSYIHLASEQAPTGVGEAALLHYVDPTGRTLGE----- 472
516 QQHIIIGGFFSELSKVV-RTYIRERVVDQLAHI-DVQLAQSVADNLGITLTDQRNLAPPK 573
473 ----FKAYPDGFLTCVPNGASSG 491
574 EVNGVKKDPSELSLYAVEGGSIKG 596

RESULT 10
US-10-132-134-34
; Sequence 34, Application US/10132134
; Publication No. US20030171562A1
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Yang, Xianshu
; APPLICANT: Staiffa, Alfredo
; APPLICANT: Zazopoulos, Emmanuel
; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES
; FILE REFERENCE: 3012-2US
; CURRENT APPLICATION NUMBER: US/10/132,134
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 8360
; TYPE: PRT
; ORGANISM: Streptomyces antibioticus
US-10-132-134-34

Query Match 4.2%; Score 118.5; DB 14; Length 8360;
Best Local Similarity 23.2%; Pred. No. 4.8;
Matches 143; Conservative 56; Mismatches 241; Indels 177; Gaps 36;
13 DGASGAGQLVPEV-----NASDPLAM-----DPVAGSGSTAVATAGQVNPIDPWI 56
1388 DGTGAARFVDPADVPAREADDPVAVVGMARYGDAEDPRALMDRLLAGDGLVEFVTRWD 1447
57 INNFVQAPQGEFT-----ISPNPTPGGVFLDLS--LGPHLNP-FLHLHLSQMN-----GW 103
1448 LGPEVTCRAGSFVRGMDRFD-----VFFAISGVEAAHMDPQORI-FLEQCNWNALEDAGY 1501
104 VGNMRVRIMLAGNAFTAGKIIVSCIPGFGSHNLT-----AQATLPHVIADVRLDPI 158
1502 TGE-RLRERNCG-----VYVGYCA--GDYDSIGSDRAPQA-LWGTGSGVVASRIAY 1549
159 EVPLEDVNRVLFHNDRNQOTMRLVCLMYTLPTRTGGGTGDSFVAVAGRVMTCPSPDFNLF 218
1550 QDLKGPALTDTTSCSSLSVSLHLAC---RDLRTGAA---DWAIAGGV-----FLQ 1594
219 LVPPTVEQKTRFTPLNPLSLSNSRAP--LPTSGMI-----SPDNVQSVQF 265
1595 TTPRLYEATRAGMLSPTRCHSFDSRADGFPVGEAGAVVLRKLSDALRQGDHVGVLVR 1654
266 QNGRCITLQRLVG-TTPVSLSHVAKIRGTSNGTV---INLTELQGTTPH---PFE----- 313
1655 ATG-VNQDGTNGTITAPSAQSQALLREVHAGVFGVQLVEAHGTGTQLGDPDIFRALS 1713

Query Match 4.1%; Score 116; DB 12; Length 2468;
Best Local Similarity 22.2%; Pred. No. 1.3;

Matches	107;	Conservative	41;	Mismatches	170;	Indels	164;	Gaps	23;
Qy	23	PEVNADPLAMPDVAGSS--TAVATAGQVNPIDPWIIINNFVQAPGEGTISFNN-----TPG	77						
Db	426	PQVNASNGSVLSGTAEGAVTIVITDGNPN-----GQTSADAGNWSFTPG	472						
Qy	78	GVFLFDLSGLGPHLNPFLHLQSMYNGWGVNMRVRIMLAGNAFTAGKIIVSCIPPGF-----	132						
Db	473	-----SOLPDGTVVNVVAR--DAAGNSSPATSIITVDGVAPNAPVVEP	512						
Qy	133	-----GSHNLITIAQATLFPHVIAADVRLDPIEVPLED--VRNVLFHNND	174						
Db	513	SNSELSGTAEPGSSVTLTDGNGNPIGQTTADANGNSFTP-STPLPDGTVVNVVARDAA	571						
Qy	175	RNQQ-----TWRLVCMLYTPLRTCGGT-----GDSFVVA-----GRVMTCPSPDF	214						
Db	572	GNSSPPASVTVDVAVAPATPVDPSPNGTTLTGTAEPGSSVTLTDGNGNPIGQVTADGSGNW	631						
Qy	215	NFLFLVPPTVEQKTRPTFLPNLPLSSLSNGRAPLPISGMGISPDNVOSVQFNGRCRTLDG	274						
Db	632	TFTPTSTP-----LPNGTVVNATATDP--SGNASSPASV-----TVDA	666						
Qy	275	RLVGTTPVSLSHVAKIRGTSN-GRVINLTLDGTP-----FHPFEGAPAI GF	320						
Db	667	VAPATPVVNPSPNGTTLTGTAEPGATVTLTDGNGNPIGQVTADGSGNWSFTP-TTPLPNGT	725						
Qy	321	PDLGGCDWHINMTQFGHSSQTYDVTPTDFVPHLSIQ-ANGI-----	364						
Db	726	V-----VNATATDASGNTSAGSSVTVDSVAPATPVINPSNNTTLTGTAEPGSSVTLT	777						
Qy	365	-GSGNYIGVL-----SWVSPPHSPSGSQVDLMKIPNYGSSITETHLAPSVYPGFGEV	417						
Db	778	DENGNPICQVTADGSGNWSFTPTPLAD-----GTVINATATDPAGNTSGGOST	826						
Qy	418	LV	419						
Db	827	TV	828						

```

RESULT 12
US-10-246-330-4
; Sequence 4, Application US/10246330
; Publication No. US20030166030A1
; GENERAL INFORMATION:
; APPLICANT: O'Toole, George A.
; APPLICANT: Mah, Thien-Pah
; TITLE OF INVENTION: METHODS TO STUDY AND MECHANISMS OF
; TITLE OF INVENTION: BIOPILM-SPECIFIC ANTIBIOTIC RESISTANCE
; FILE REFERENCE: 14537-002001
; CURRENT APPLICATION NUMBER: US/10/246,330
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 60/323,241
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 2468
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-10-246-330-4

```

```

Query Match      4.1%; Score 116; DB 14; Length 2468;
Best Local Similarity 22.2%; Pred. No. 1.3;
Matches 107; Conservative 41; Mismatches 170; Indels 164; Gaps 23;

Qy 23 PEVNASPLAMDVPVAGSS-TAVATAGQVNPIDPWIINNFVQAPQGETISNN-----TPG 77
   |||||: : : : : ||||| : : : : : |||
Db 426 PQVNASGVSLSGTAEAGVTITDGNPIT-----GQTSADANGNWSFTPG 472

Qy 78 GVLFDLSLGPHLNDPLFLHLQSMYNGVGNKRVIRIMLAGNAFTAGKIIVSCIPGF-----132
   ||||: : : : : ||||| : : : : : |||
Db 473 -----SOLPDGTVVNVVAR--DAAGNSPATSITVDGVARNAPNVPEP 512

```

Qy	133	-----GSHNLTIAQATLPHFVIADV---	TLDPIEVPLED--	VRNVLPHHND	174
Db	513	SNGSELGCTABPGSSVTLTDGNGNPIGOTTADANGNMSFTP-STPLPDGTVVNVVARDAA	571		
Qy	175	RNOQ-----TWRLVCMLYTLPLRTGGGT-----	GDSFVVA-----	GRVMTCPSPDF	214
Db	572	GNSFPPASVTVDAVAPATPTVDPNSGNTLGTABPGSSVTLTDGNGNPIGOTTADGSGNW	631		
Qy	215	NFLFLVPPTVEQKTRPFTLPNLPLSSLSNSRAPLPIGCMGISPDNVQSVQFQNGRCTLDG	274		
Db	632	TFTFSTP-----LPNGTVVNATADP-SGNASSPASV-----	TVDA	666	
Qy	275	RLVCTTPVSLSHVAKIRGTSN-GTVINLTDLGCTP-----	FHPPEGPAPIGF	320	
Db	667	VAPATPVVNPNSGNTLGTABPGATVTLTDGNGNPIGOTTADGSGNMSFTP-TTLPNPGT	725		
Qy	321	PDLGCDWHINMTQGHSSOTQYDVTDPDFVPHLSIQ-ANGI-----	-----	364	
Db	726	V-----VNATADASGNTSAGSSVTVDSVAPATPVINPSNGNTLGTABPGSSVTLT	777		
Qy	365	-GSGNYIGVL-----	SNVSPPHSPSGSQVDLWKIPNYGSSITEATHLAPSVYPPGFGEV	417	
Db	778	DGNGNPIGOTTADGSGNMSFTPSTPLAD-----	-----	GTVVNATADPAGNTSGQGST	826
Qy	418	LV	419		
Db	827	TV	828		

```

RESULT 13
US-10-157-031-171
; Sequence 171, Application US/10157031
; Publication No. US20030108890A1
; GENERAL INFORMATION:
; APPLICANT: Baranova, A. V.
; APPLICANT: Yanovsky, N. K.
; APPLICANT: Kozlov, A. P.
; APPLICANT: Lobashev, A. V.
; APPLICANT: Krukovskaya, L. L.
; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
; FILE REFERENCE: 2760-103
; CURRENT APPLICATION NUMBER: US/10/157,031
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 415
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 171
; LENGTH: 878
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-157-031-171

```

Query Match	4.0%;	Score 112.5;	DB 14;	Length 878;
Best Local Similarity	19.9%;	Pred. No. 0.55;		
Matches	90;	Conservative 42;	Mismatches 150;	Indels 171; Gaps 19
Qy	187	YTPRLRTGGGTGDSFWAGRVMTCPS-----PDENFLFLVPPTVQKTRPPTLPNLP	237	
Db	17	PTTAETG-----VTSTPSSPSLSLTDPTTSLRLTLPUSLSTSTSLTTTTDLP	64	
Qy	238	LSSLSNRAPLPISGMGI SPNVQSVQFQNGRCTL DGRLVGTPPTVSLSHV-AKIRGTSNG	296	
Db	65	SIPTDISLPTPIHIISSSP-SIQSTETSS-----LVGTTSPMTSVTRATLASTENT	115	
Qy	297	TVIN-----LTELDGTGFPHFEGGAPATGFPDPLGGCDWH-----	329	
Db	116	PISSEFSTSVVTPETPTTQAPPVLMSATGQTSPVPTTVTFGSDMSSTSLTLTLPSTAL	175	
Qy	330	---LNMTQFGHSSQTQVDVDTPTDFVPHL-----GSIQANGIGSGNY	369	
Db	176	SKIMSTSQFPPIPSHTSSLTQTPPS--IPSLQTSLTSTSEPTTESFTRGSTSNALTS---	231	
Qy	370	IGVLSWVSPP-----SHPGSGQV-----DLMKIPINYGSS	398	

```
Db 232 ESTIWSSTPTIIMSSSPSSASITVPFATTIHSVPSSPYIFSTENVGSASITAPFSLSS 291
QY 399 ITEATHLAPSVVPPGFEVLVPE----FMSKIPGCGAYSLPCLLPQOEYISHLASEQAPTVG 454
Db 292 STTSTSPSSSTATTALTEITPESYISLSTTPCPGTITI-TIVP-----ASTPDCV- 342
QY 455 EAALLHYVDPDGTGR-----LGEFKAYPDGGLT 482
Db 343 -----EMDPSTEATSPPTPLTVPPFTTETMVTCPSSISMOTTLATHMDTSSMTPESSS 396
QY 483 CVPNGASS-GPOOLPINGVVFV-----SWVS 508
Db 397 IIPNASSSTGTGTVPNTVFTSTRLPTSETWLS 429

RESULT 14
US-10-087-192-366
; Sequence 366, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 366
; LENGTH: 702
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-366

Query Match 4.0%; Score 112; DB 12; Length 702;
Best Local Similarity 23.7%; Pred. No. 0.43;
Matches 75; Conservative 41; Mismatches 114; Indels 86; Gaps 16;

QY 210 PSPDFNLEFL-VPTTVQKTR-PTTLNPLSLSSNSRAPLPTSGMGISPDNVQSVQFQN 267
Db 290 PSNGVEFISVQSPPTVSGLTKNPVSLFSLPNPTKPNKVPSPSPSIQGNP----- 339
QY 268 GRCTLDGLVCTT-----PVSLSHVAKIRGTSNGTVINLTDLGTDPFHPFEGPAPIGFP 321
Db 340 ---TASAAPLTGTTLAVQAVPTAHIVQATR-TSLPTV-----GPSGLYSP 380
QY 322 --DLGGCDWHINMTQFGHSSQTVQDVTPTTFVPHLGSIQANGIGSGNVIG-VLSVW-- 376
Db 381 STNRGPIQMKIPISAFSTSSAAEQNSNTTPRIDPKLNHTPVMTSSSQVPSRLQPIQ 440
QY 377 SPPHPSG-----SQVDLWKIPNYGSS-----ITEATHLAPSVVPPGFEVLVPEFMS 423
Db 441 APPLQPSGVPTSGPSQTTIHLPTAPTIVNTHRPVQTVTLRPVPAPANHQV-----YT 497
QY 424 KIPGGAYSLPCLLPQOEYISHLASEQAP---TVGEAALLHYVDPDGTGRTLGEFKAYPDGF 480
Db 498 TLPAEPA-----QAPLRGTVMQAPAVRQVFNQNSVTV-----RVPQTT 535
QY 481 LTCVPNG---ASSGPQ 493
Db 536 TYVNNGLTLGSGTGPQ 551
```

RESULT 15
US-10-424-599-212925
; Sequence 212925, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

```
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 212925
; LENGTH: 769
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(769)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_34297C.1.pep
US-10-424-599-212925
```

```
Query Match 4.0%; Score 112; DB 12; Length 769;
Best Local Similarity 21.3%; Pred. No. 0.5;
Matches 73; Conservative 43; Mismatches 116; Indels 110; Gaps 17;

QY 232 TLPNPLSLSSNSRAPLPI---SCMGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVA 288
Db 356 TFSEMPSSSQTGAEYPAHTAFSNGMGIGLSNL-----TMDG-----METDETR 398
QY 289 KIRGTSNGTVINLTDLGT-----PFHPFEGPAPIGFPDLGGCDWHI 330
Db 399 PAEGSQHGLNTNTYSLNGLMLGLSRQTANRGVLSSEFQGFHP-----PFSRDPGWEI 451
QY 331 NMTQFGHSSQTVQDVTPTTFVPHLGSIQANGIGSGNVIGVLSWVSPSSH----- 381
Db 452 PFLHGWINGQSQVGV-----PSMLPHMGASRDN---LSQHIGSSSYIKASNHSSTNVDAAM 504
QY 382 -----PSGSQVDLWK-----IPNYGSS-----ITEATHLAPSVVPPGFEVLVPEFMSKIPGP 428
Db 505 PSSAIPGGISITQITSMRSGLENHFSQSRVPVSESGNLAASINTPHDGFDMQTVIVSRIQSE 564
QY 429 GAYS-----LPCLLPQOEYISH-LASEQAPTVGEAALLHYVDPDGTGRTLGEFKAYPDGF 480
Db 565 LATSVAATAAELPCTVKLRVMSHDIKNPCAP-----LNAADP-----GL 603
QY 481 LTCVPN---GASSGPOOLPINGVVFVSVWSRFYQLKPVGTAS 520
Db 604 QTPVHQDPGVATSPTRHPISAHQV---WDAR--MLPPLEEAT 640
```

Search completed: June 1, 2004, 14:04:29
Job time : 36.8809 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:36:02 ; Search time 10.285 Seconds
(without alignments)
4985.230 Million cell updates/sec

Title: US-09-926-799-2
Perfect score: 2819
Sequence: 1 MMASKDATSSVDGASGAGQ.....YQLKPVGTASSARGRLGLRR 530

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:
1: pir1:
2: pir2:
3: pir3:
4: pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2789	98.9	530	2 B37471	capsid protein - N
2	2033.5	72.1	546	2 B37491	major capsid prote
3	1238	43.9	542	2 S60616	capsid protein - h
4	1148.5	40.7	539	2 S40111	capsid protein - h
5	366.5	13.0	576	2 A53982	capsid protein - E
6	326	11.6	2344	2 S53399	genome polyprotein
7	321.5	11.4	2344	2 S64740	genome polyprotein
8	319.5	11.3	2344	1 RWRWRH	genome polyprotein
9	261.5	9.3	702	1 A48562	coat protein - San
10	251	8.9	703	1 C48562	coat protein - San
11	242.5	8.6	668	2 JQ2354	capsid protein - f
12	241.5	8.6	668	1 VCMWFC	coat protein - fel
13	240.5	8.5	668	1 VCMWFF	coat protein - fel
14	237	8.4	671	1 VCMWF9	coat protein - fel
15	229.5	8.1	668	2 JQ2356	capsid protein - f
16	149	5.3	2205	1 GNNY2W	genome polyprotein
17	147	5.2	2207	1 GNNY5P	genome polyprotein
18	142	5.0	2194	1 GNNY27	genome polyprotein
19	140.5	5.0	2206	1 GNNY27	genome polyprotein
20	136.5	4.8	2185	1 GNNYSV	genome polyprotein
21	135	4.8	2508	2 S61441	surface-associated
22	134.5	4.8	2164	1 GNNY89	genome polyprotein
23	132.5	4.7	2207	2 S09553	genome polyprotein
24	129	4.6	3473	1 A46112	genome polyprotein
25	129	4.6	3473	2 S27927	polyprotein - rice
26	128.5	4.6	2185	1 GNNYSH	genome polyprotein
27	127.5	4.5	2175	1 GNNYBE	genome polyprotein
28	126	4.5	2185	1 JQ2021	genome polyprotein
29	126	4.5	2206	1 GNNY21	genome polyprotein

30	126	4.5	2207	1 GNNY1P	genome polyprotein
31	125	4.4	3085	2 T00327	polyprotein - infe
32	124	4.4	2206	2 S03822	genome polyprotein
33	123.5	4.4	929	2 A44048	genome polyprotein
34	123.5	4.4	2183	2 T42764	coagulation factor
35	123	4.4	1200	2 T17404	hyalin - sea urchi
36	122	4.3	2206	1 GNNY4P	genome polyprotein
37	122	4.3	2209	1 GNNY2P	genome polyprotein
38	121.5	4.3	799	2 S75351	hypothetical prote
39	120	4.3	477	1 P2WL42	L2 protein - human
40	119.5	4.2	2214	1 A48548	genome polyprotein
41	119.5	4.2	3716	2 E70969	probable ppe prote
42	117.5	4.2	1436	2 B70520	probable ppe prote
43	117	4.2	2179	1 GNNYH4	genome polyprotein
44	116	4.1	2209	1 GNNY3P	genome polyprotein
45	116	4.1	2468	2 A83412	hypothetical prote

ALIGNMENTS

RESULT 1

B37471 capsid protein - Norwalk virus
C:Species: Norwalk virus
C:Date: 24-Feb-1994 #sequence revision 25-Apr-1997 #text_change 28-Jul-2000
C:Accession: B37471
B37471 X. Wang, M. Wang, K. A. Estes, M. R. Virology 1996, 50:617-693
A:Title: Sequence and genomic organization of Norwalk virus.
A:Status: preliminary; not compared with conceptual translation
A:Reference number: A37471; MUID:93303939; PMID:8391187
A:Accession: B37471
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: genomic RNA
A:Residues: 1-530 <VIA>
A:Cross-references: GB:M87661; NID:gl061311; PIDN:AAB50466.1; PID:gl061313
A:Note: sequence extracted from NCBI backbone (NCBIP:134157)
C:Superfamily: human calicivirus capsid protein

Query Match	98.9%;	Score	2789;	DB 2;	Length	530;			
Best Local Similarity	98.5%;	Pred. No.	1.2e-189;						
Matches	522;	Conservative	5;	Mismatches	3;	Indels	0;	Gaps	0;
QY	1	MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIINNPF	60						
DB	1	MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIINNPF	60						
QY	61	VOAPOGEFTISPNPTGGVLFDSLGLPHLNPFLLHLSQMYNGWGNMVRIMLAGNAFTA	120						
DB	61	VOAPOGEFTISPNPTGCVLFDLSLGLPHLNPFLLHLSQMYNGWGNMVRIMLAGNAFTA	120						
QY	121	GKIIIVSCIPPGSGSHNLTAQATLFPHVIAVRTLDPIEVLDPEDVRNVLPHNNDRNQOTM	180						
DB	121	GKIIIVSCIPPGSGSHNLTAQATLFPHVIAVRTLDPIEVLDPEDVRNVLPHNNDRNQOTM	180						
QY	181	RLVCMLYTPLRTGGGTGDSFVAVAGRVMTCPSPDFNPLFLVPPTVEQKTRPFTLPNPLSS	240						
DB	181	RLVCMLYTPLRTGGGTGDSFVAVAGRVMTCPSPDFNPLFLVPPTVEQKTRPFTLPNPLSS	240						
QY	241	LSNSRAPLPISGMGISPDNVQSVQFNGRCTLDGLRVGTTPVLSLSHVAKIRGTSNGTVIN	300						
DB	241	LSNSRAPLPISGIGISPDNVQSVQFNGRCTLDGLRVGTTPVLSLSHVAKIRGTSNGTVIN	300						
QY	301	LTELDTGTFPHPEGPAPIGFDPDGGCDWHNNMTQFGHSSQTOYDVTDTTDTFVPHLGSIQ	360						
DB	301	LTELDTGTFPHPEGPAPIGFDPDGGCDWHNNMTQFGHSSQTOYDVTDTTDTFVPHLGSIQ	360						
QY	361	ANGIGSGNYIGVLSWSPSPHSGSODLWKIPNYGSSITEATHLAPSVYPGFGELVLF	420						
DB	361	ANGIGSGNYGVLSWISPSHSGSODLWKIPNYGSSITEATHLAPSVYPGFGELVLF	420						
QY	421	FMSKIIPGCGAYSLPCLLPQEIYISHLASEQAPTVGEAALLHYVDPTGRTLGEFKAYPDGF	480						

Db 421 FMSKMPGAYNLPCLLPQYIYSHLSEQAPTVGEAALLHYVDPDTGRNLGEFKAYPDGF 480
 QY 481 LTCVNGASSGQQLPQYIYSHLSEQAPTVGEAALLHYVDPDTGRNLGEFKAYPDGF 530
 Db 481 LTCVNGASSGQQLPQYIYSHLSEQAPTVGEAALLHYVDPDTGRNLGEFKAYPDGF 530

RESULT 2
 B37491
 major capsid protein [similarity] - Southampton virus
 N/Alternate names: orf2 protein
 C/Species: Southampton virus
 C/Date: 22-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 18-Aug-2000
 C/Accession: B37491
 R/Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
 Science 259, 516-519, 1993
 A/Title: Sequence and genome organization of a human small round-structured (Norwalk-like) virus
 A/Reference number: A37491; MUID:93142023; PMID:8380940
 A/Accession: B37491
 A/Status: not compared with conceptual translation
 A/Molecule type: nucleic acid
 A/Residues: 1-546 <LAM>
 A/Cross-references: GB:L07418; NID:gl236787; PIDN:AAA92984.1; PID:g295114
 A/Note: sequence extracted from NCBI backbone (NCBIP.123458)
 C/Species: small round-structured virus, SRSV, Norwalk virus, Norwalk-like virus, serotype
 C/Superfamily: human calicivirus capsid protein
 C/Keywords: Glycoprotein
 F/303,340,441/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 72.1%; Score 2033.5; DB 2; Length 546;
 Best Local Similarity 69.6%; Pred. No. 3.5e-136; Mismatches 82; Indels 17; Gaps 6;
 Matches 380; Conservative 67;
 QY 1 MMASKDQTSVDCGAGQGLVPEVNASDPLAMDVPVAGSSTAVATAGQVNPIDPWIINF 60
 Db 1 MMASKDQTSVDCGAGQGLVPEVNASDPLAMDVPVAGSSTAVATAGQVNPIDPWIINF 60
 QY 61 VQAPQGEFTTSPNTPGGVLDLGLPHLNPFLHLSQMYNGVGNVRVIMLAGNAFTA 120
 Db 61 VQAPQGEFTTSPNTPGGVLDLGLPHLNPFLHLSQMYNGVGNVRVIMLAGNAFTA 120
 QY 121 GKIVSCIPPGFGSHNLTIAQATLPHVIAVDRVLTDPVLEPVDVRLVFNHNDNQOTM 180
 Db 121 GKIVSCIPPGFGSHNLTIAQATLPHVIAVDRVLTDPVLEPVDVRLVFNHNDNQOTM 180
 QY 121 GKIVCCVPPGFTSSSLTIAQATLPHVIAVDRVLTDPVLEPVDVRLVFNHNDNQOTM 179
 Db 121 GKIVCCVPPGFTSSSLTIAQATLPHVIAVDRVLTDPVLEPVDVRLVFNHNDNQOTM 179
 QY 181 RLVCMLYPLRTGGTG--DSFVAVGRVMTCPSPDFNLFVLPVPTVEQKTRPFLPNLPL 238
 Db 181 RLVCMLYPLRTGGTG--DSFVAVGRVMTCPSPDFNLFVLPVPTVEQKTRPFLPNLPL 238
 QY 239 SSLNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN--G 296
 Db 239 SSLNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN--G 296
 QY 240 QTLNSRPFSLIQGMILSPDASQVQVQNGRCLDGLGTLTPATSGQLFVRGKINQGA 299
 Db 240 QTLNSRPFSLIQGMILSPDASQVQVQNGRCLDGLGTLTPATSGQLFVRGKINQGA 299
 QY 297 TVINLTDLGTPFHFEGPAPIGFDPDLGGCDWHINMTQGHSS-----QTOYVDTPDPT 351
 Db 297 TVINLTDLGTPFHFEGPAPIGFDPDLGGCDWHINMTQGHSS-----QTOYVDTPDPT 351
 QY 300 RTLNLTEDVGRKPMADFAPVGFDPFGKCDWHIRISKTPNNTGSGDPMRSVSQVNTVQ 359
 Db 300 RTLNLTEDVGRKPMADFAPVGFDPFGKCDWHIRISKTPNNTGSGDPMRSVSQVNTVQ 359
 QY 352 FVPHLGSIQANGI---GSGNYIGVLSWSPSPHSGSQVDLWKIPNYGSSITEATHLAPS 408
 Db 352 FVPHLGSIQANGI---GSGNYIGVLSWSPSPHSGSQVDLWKIPNYGSSITEATHLAPS 408
 QY 360 FVPHLGSIQANGI---GSGNYIGVLSWSPSPHSGSQVDLWKIPNYGSSITEATHLAPS 419
 Db 360 FVPHLGSIQANGI---GSGNYIGVLSWSPSPHSGSQVDLWKIPNYGSSITEATHLAPS 419
 QY 409 VYPPGEGEVLVFMKIPGPGAYS---LPCLLPQYIYSHLSEQAPTVGEAALLHYVDP 464
 Db 409 VYPPGEGEVLVFMKIPGPGAYS---LPCLLPQYIYSHLSEQAPTVGEAALLHYVDP 464
 QY 420 VFPFGEGEALVYFSAFPPGNRRSAPNDVPCLLPQYIYTHFVSEQAPTMGDAALLHYVDP 479
 Db 420 VFPFGEGEALVYFSAFPPGNRRSAPNDVPCLLPQYIYTHFVSEQAPTMGDAALLHYVDP 479
 QY 465 DTGRTLGEFKAYPDGFLTCVNGASSGQQLPQYIYSHLSEQAPTVGEAALLHYVDP 524
 Db 465 DTGRTLGEFKAYPDGFLTCVNGASSGQQLPQYIYSHLSEQAPTVGEAALLHYVDP 524
 QY 480 DINRNLGEFKAYPDGFLTCVNGASSGQQLPQYIYSHLSEQAPTVGEAALLHYVDP 539
 Db 480 DINRNLGEFKAYPDGFLTCVNGASSGQQLPQYIYSHLSEQAPTVGEAALLHYVDP 539
 QY 525 RLGLRR 530
 Db 540 RLGLRR 545

RESULT 3
 S60616
 capsid protein - human calicivirus (strain Melksham)
 C/Species: human calicivirus
 A/Variety: strain Melksham
 C/Date: 23-May-1997 #sequence_revision 23-May-1997 #text_change 28-Jul-2000
 C/Accession: S60616
 R/Green, S.M.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
 Virus Res. 37, 271-283, 1995
 A/Title: Capsid diversity in small round-structured viruses: molecular characterization
 A/Reference number: S60615; MUID:96136658; PMID:8533462
 A/Accession: S60616
 A/Molecule type: genomic RNA
 A/Residues: 1-542 <GRE>
 A/Cross-references: EMBL:X01879; NID:g976077; PIDN:CAAS7462.1; PID:g976079
 A/Experimental source: strain Melksham
 A/Note: it is uncertain whether Met-1 or Met-3 is the initiator
 C/Superfamily: human calicivirus capsid protein
 C/Keywords: capsid protein; coat protein
 Query Match 43.9%; Score 1238; DB 2; Length 542;
 Best Local Similarity 47.5%; Pred. No. 6.6e-80;
 Matches 264; Conservative 74; Mismatches 168; Indels 50; Gaps 13;
 QY 1 MMASKDQTSVDCGAGQGLVPEVNASDPLAMDVPVAGSSTAVATAGQVNPIDPWIINF 60
 Db 1 MMASKDQTSVDCGAGQGLVPEVNASDPLAMDVPVAGSSTAVATAGQVNPIDPWIINF 60
 QY 61 VQAPQGEFTTSPNTPGGVLDLGLPHLNPFLHLSQMYNGVGNVRVIMLAGNAFTA 120
 Db 61 VQAPQGEFTTSPNTPGGVLDLGLPHLNPFLHLSQMYNGVGNVRVIMLAGNAFTA 120
 QY 121 GKIVSCIPPGFGSHNLTIAQATLPHVIAVDRVLTDPVLEPVDVRLVFNHNDNQOTM 180
 Db 121 GKIVSCIPPGFGSHNLTIAQATLPHVIAVDRVLTDPVLEPVDVRLVFNHNDNQOTM 180
 QY 176 RLVCMLYPLRTGGTG--DSFVAVGRVMTCPSPDFNLFVLPVPTVEQKTRPFLPNLPL 176
 Db 176 RLVCMLYPLRTGGTG--DSFVAVGRVMTCPSPDFNLFVLPVPTVEQKTRPFLPNLPL 176
 QY 181 RLVCMLYPLRTGGTG--DSFVAVGRVMTCPSPDFNLFVLPVPTVEQKTRPFLPNLPL 240
 Db 181 RLVCMLYPLRTGGTG--DSFVAVGRVMTCPSPDFNLFVLPVPTVEQKTRPFLPNLPL 240
 QY 236 RLVCMLYPLRTGGTG--DSFVAVGRVMTCPSPDFNLFVLPVPTVEQKTRPFLPNLPL 236
 Db 236 RLVCMLYPLRTGGTG--DSFVAVGRVMTCPSPDFNLFVLPVPTVEQKTRPFLPNLPL 236
 QY 241 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 298
 Db 241 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 298
 QY 296 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 296
 Db 296 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 296
 QY 346 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 346
 Db 346 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 346
 QY 356 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 356
 Db 356 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 356
 QY 390 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 390
 Db 390 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 390
 QY 405 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 405
 Db 405 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 405
 QY 447 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 447
 Db 447 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 447
 QY 465 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 465
 Db 465 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 465
 QY 507 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 507
 Db 507 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 507
 QY 523 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 523
 Db 523 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 523
 QY 539 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 539
 Db 539 LNSRAPLPISGMIGISPDNVQSVQFONGRCTLDGRLVGTTPVLSLHVAKIRGTSN-- 539
 RESULT 4
 S40111
 capsid protein - human calicivirus (strain Bristol isolate B493)
 C/Species: human calicivirus
 A/Variety: strain Bristol isolate B493
 C/Date: 25-Dec-1994 #sequence_revision 27-Feb-1997 #text_change 28-Jul-2000
 C/Accession: S40111

R; Green, S.M.; Dingle, K.E.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N. submitted to the EMBL Data Library, December 1993
A; Description: Human enteric caliciviridae: a new prevalent SRSV group defined by RNA-de
A; Reference number: S40111
A; Accession: S40111
A; Molecule type: genomic RNA
A; Residues: 1-539 <GRE>
A; Cross-references: EMBU:X76716; MID:G436410; PIDN:CAA54134.1; PID:G436411
A; Experimental source: human enteric calicivirus strain Bristol isolate B493
C; Superfamily: human calicivirus capsid protein
C; Keywords: capsid protein; coat protein

Query Match	40.7%	Score 1148.5;	DB 2;	Length 539;
Best Local Similarity	45.1%;	Pred. No. 1.4e-73;		
Matches	249;	Conservative	82;	Mismatches 182;
			Indels	39;
			Gaps	14;
QY	1	MMMASKDAYTSSVDGASGAGOLVPEVNASDPLAMPDVAGSSTAVATAGQVNPIDPWIINN	60	
DB	1	MKWASNDANPS-DGS--AANLVPEVN-NEWMALEFPVVGAAIAAPVAGQONVIDPWIRN	56	
QY	61	VQAPQGEFTISPNNTPGGVLFDLSLGHLPNPFLLHLGSLQMYNGWGVNMVRIMLAGNAFTA	120	
DB	57	VQAPGGEFTVSPRNAPGEILWSAPLGPDLNPNVLSHLRMYNGYAGGFEVQVILAGNAFTA	116	
QY	121	GKLIIVSCIPGFGSHNLTAQATLFEHVITADVRLDPIEVLPELDVENVLFHNDRNQOTM	180	
DB	117	GKVFIAAAPPNFTEGLSPSQVTMFPHIIVDRVQLSEPVLIPLPDVRNNFYHYNQANDSTL	176	
QY	181	RLVCMLYTPLRITCGGTGDSFVAVRGVMTCPSPDNFELFLVPPTVEOKTFPTLPNLPSS	240	
DB	177	KLIAMLYTPLRANNAGDDVFTVSCRVLTPSPDPDFILFVPTTVESRTKFTFVPLTVBE	236	
QY	241	LSNSRAPLPISGMGIISPNNVQSVOFGNGRCTLDGRVLVGTTPVSLSHVAKIRGT-----S	294	
DB	237	MSNSRFPPIPLEKLYTGPFSAFVQPGNGRCTTDGVLGTTQLSAVNINCNRFGDVTHIAGS	296	
QY	295	NGTVINILTDGTPRPFEG-PAPIGFDPDLGGCDWHNNMQF-----GHSSQTYQDVDTTP	349	
DB	297	HDYTMNLASQNSWNYDPTTEIPAPLGPDPVFGKIQGL-LTQTTADGSTRAHKATVSTGS	355	
QY	350	DTFVPHLGSIQ-----ANGIGSGNY-----IGVLSWVSPFSPGSDVLWKIPNYGSSI	399	
DB	356	VHFTPKLGSVQFTDTNDNFQAGQNTKFTPVGVIQ--DGDHHQNEPQ--QWLLPNVSGRT	411	
QY	400	TEATHLAPSVYPGCGEVLVFFMSKIPGCGAY----SLPCLLPQEIYISHLASQAPVTGBA	456	
DB	412	GHNVHLAPAVAPTTPGGEQLLFFRSTMPGCGSGYPNNMLDCLLPQEWVHLHFYQEAAPQSDV	471	
QY	457	ALLHYVDPTDGTGRLTGFFKAYPDGFLTCVPNGASSGPOQ--LPIINGVFVSVMSVRFYQLK	514	
DB	472	ALLRFVNDPTGTVLFPCKLHKSQYITV-----AHTGPYDLVLPNGVFRFDSWNGFYTLA	527	
QY	515	PVGTASSARGRL	526	
DB	528	PMNGTGRRRAL	539	

RESULT 5
AS39982
capsid protein - European brown hare syndrome virus
C/Species: European brown hare syndrome virus
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 28-Jul-2000
C/Accession: AS39982
R./Wirblich, C.; Meyers, G.; Ohlinger, V.F.; Capucci, L.; Eskens, U.; Haas, B.; Thiel, H.
J. Virol. 68, 5164-5173, 1994
A./title: European brown hare syndrome virus: relationship to rabbit hemorrhagic disease
A./Reference number: AS3982; MUID:94309183; PMID:7518531
A./Accession: AS3982
A./Status: preliminary
A./Molecule type: mRNA
A./Residues: 1-576 <WIR>
A./Cross-references: GB:U09199
C./Superfamily: human calicivirus capsid protein

```

Query Match      13.0%; Score 366.5; DB 2; Length 576;
Best Local Similarity 26.0%; Pred. No. 3.3e-18;
Matches 137; Conservative 81; Mismatches 221; Indels 87; Gaps 21;

QY      9  TSSVDCGASGAGQLVPEVNADPLAMPDVAGSSATAVATAG-----QVNPDIPWLNINPVQ 62
      15  TASVPGTTTGDGMDPGVAVSTDVVTADNVAAS-----VATAGIGGGPQQASPOESWRVNFY- 70
QY      63  APOGEFTTIPNNTPGGVLFDSLGLPHLNPFLHLSQMYNGVMGNMRYRIMLAGNAFTAGK 122
      71  --NDVFTWSVTDAPGSSILYSVQHSPPQNNPTQLVSQMYACWAGCMQPRFFIVAGSGIEGGR 128
QY      123  IIVSCITPPGEG-SHNLTIAQAATLPPHVIADVRITLDPIEVLPLEDVNRVLFHNDRNRQQTMR 181
      129  LVCAIITPGGIQIQGLEVRQ---FPHVVIDARSLPEVITIIMPDLRPEMYHPTGDPGLVPT 188
QY      182  LVCMLYTPI-RTGGGTGDSFVAGRVYMTCPSPDNFLFLVPPTVE---OKTRPFTPLNPL 238
      186  LVVSVVNNLINPGGTTSAIQVT--VETRESEDFEFLIRAPASKTVDSVNPSSLTTPV 243
QY      239  --SSLSNSRAPLPIFGMGIISPDNVQSVQFQNGRCTLDRLVGTTPVSL-----SHVAK 289
      244  LTCAGSDNRWGAIVGLQVPVGGFST---SNRHNWNGSTYGVSSPRFDDIDHPSGNVSY 300
QY      290  IRTSGNTVINLTTELQGT-PPHPFEGPAPIGPDLGCGDWH-----INMTQFGHSS 339
      301  PTGSATNTTETVYANAGTATTNPISINAPDGFDPDMAIPESGTTIPTGAWVGQGVVNAS 360
QY      340  QTOYDVTTPDTFVPHLGSIQANGIGSGNYIGVLSWVSPSPHSPSGSOVDLWKIPNYGSSI 399
      361  NG-----TPVGVTVQAVELGFAN--GAPSSIRPVTTTGTGAQLVAKSI--YGVAI 405
QY      400  TEATHLAPSVYPGFGVELVFFNMKSITPGGAYSLPCLL-----POYIISHLSEQAPT 452
      406  AQNOSSAG-----IIFLSK---GMVSTPGVAATYTPQPSAIVTTPGTPVAAP 450
QY      453  VGEAALLHYVDDPTGRTLGEFKAYPDGFLTCVPNGASSGGPQQLPIN 498
      451  IGNKTPIMF--SAVVRRTGDVNAGPGSV-----NGTQYGVGSGPFLS 489

RESULT 6
SS5399
genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)
C:Species: rabbit hemorrhagic disease virus
A:Variety: isolate BS89
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 18-Jun-1999
C:Accession: SS5399
R:Rossi, C.
submitted to the EMBL Data Library, May 1995
A:Reference number: SS5399
A:Accession: SS5399
A:Molecule type: Genomic RNA
A:Residues: 1-2344 <ROS>
A:Cross-references: EMBL:X87607; NID:G9854640; PIDN:CAA60910.1; PID:G9854641
A:Experimental source: isolate BS89
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: polyprotein

```

	Query Match	11.6%;	Score 326;	DB 2;	Length 2344;	
	Best Local Similarity	25.7%;	Pred. NO. 1.6e-14;			
	Matches 139;	Conservative 66;	Mismatches 171;	Indels 164;	Gaps 27;	
Qy	4	ASKDATSSVDGASGACQLVPEVNASDPLAMDVPAGSSTAVATAG-----QVNPIDPWII	57			
	:	: :	:	:	:	:
	:	: :	:	:	:	:
Db	1779	AGTATTASVPGTTDG-LDPGVVATTSV--TAENSASIATAGIGCGPPQOVQOETWRT	1835			
	:	:	:	:	:	:
Qy	58	NNFVQAPOQEFTISPNNTGGVLFDLSGLPHLNPILLHSOMYGVGNMRVRIMLAGNA	117			
	:	:	:	:	:	:
Db	1836	NFY----NDVFWSVADAPCSILTYVQHSPQNPFPTAVLSQMTAGWAGGQCFRIVVAGSG	1892			
	:	:	:	:	:	:
Qy	118	F7AGKLIIVSCIPPGFG-SHNLITIAQTALTPHVIADVRTLDPIEVPEDVRNVLFIHNDRN	176			
	:	:	:	:	:	:

```

1893 VFGRLVAIVPPGIEIGPLEVRQ---FPHVVIDARSLEPVTITMPDLRENMYHTGDP 1949
QY 177 QOTMRLVCMLYTPLRTGGTGDSPVAGRVMTCPSPDNELFLVPTVEQKTRPTLPNL 236
Db 1950 GLVPTLVLSVYNNLINPFG-GSTSAIQVTETRPSEDFEFVMI-----1991
QY 237 PLSSLSNSRAPLPISGMIGSPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGTS-- 294
Db 1992 -----RAPSKVDSISPAGLL-----TTPV-----LTGVND 2019
QY 295 ---NGTVINLTDLGTFPHFPGPAPIGPDLGGCD--WHINMTOFGHSSQOTQYDVTTP 349
Db 2020 NRWNGQIVGLQ-----PVBGGF---STCNRHNLNGSTYGWSSPRFADID--- 2061
QY 350 DTFVPHLGIQANGIGSGNYIVGLSWSPSPSHSGSDVLWKIPYNGSSITEATHLAPSV 409
Db 2062 -----HRRGSAST-----PGSNATNVLOFW-YANAGSAVDNP---ISQV 2096
QY 410 YPPGFEVLVFFMSKIP--GPGAYSILPCLLPQEIYSHLASEQAPTVEGALLHYVDDTG 467
Db 2097 APQGFDP-----NSFVFPNGPG--IPA-----AGWVGFAIW---NSNSG 2131
QY 468 R-TLGBEKAYPDGFLTCVP-----NGAS-----SGPQQLPINGVEVFVSVW 507
Db 2132 APNVTTVQAYELGFATGAPGNLQPTTNTSGAQTVAKSIVAVVTGTAQNPFA-GLFVNAGCV 2190

RESULT 7
S64740
genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)
N:Contains: VP60 protein
C:Species: rabbit hemorrhagic disease virus
A:Variety: isolate AST/89
C:Date: 12-Jul-1996 #sequence revision 26-Jul-1996 #text_change 20-Jun-2000
C:Accession: S64740; S4944; S49018; S65012
R:Casais, R.; Martin-Alonso, J.; Boga, J.; Parra, F.
A:Description: Genomic organization of rabbit hemorrhagic disease virus determined by di
A:Reference number: S64740
A:Accession: S64740
A:Molecule type: genomic RNA
A:Residues: 1-2344 <CNS>
A:Cross-references: EMBL:Z49271; NID:g1182032; PIDN:CAA89265.1; PID:g1150552
A:Experimental source: isolate AST/89
R:Boga, J.; Casais, R.; Martin, M.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra, F.
submitted to the EMBL Data Library, July 1993
A:Description: Molecular cloning, sequence and expression of the capsid protein gene fro
A:Reference number: S46944
A:Accession: S46944
A:Molecule type: genomic RNA
A:Residues: 1650-2344 <BOG>
A:Cross-references: EMBL:Z24757; NID:g515622; PIDN:CAA80881.1; PID:g515623
A:Experimental source: isolate AST/89
R:Parra, F.; Boga, J.A.; Martin, M.S.; Casais, R.
Virus Res. 27, 219-228, 1993
A:Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus supp
A:Reference number: S49018; MUID:93255896; PMID:8488721
A:Accession: S49018
A:Molecule type: genomic RNA
A:Residues: 1650-1796 <PAR>
A:Cross-references: EMBL:Z24757
A:Experimental source: isolate AST/89
A:Accession: S65012
A:Molecule type: protein
A:Residues: 1767-1779;1875-1877,'X',1879-1881;1936-1938,'X',1940-1941 <PAW>
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: blocked amino end; polyprotein

Query Match 11.4%; Score 321.5; DB 2; Length 2344;
Best Local Similarity 24.7%; Pred. No. 3.4e-14;
Matches 136; Conservative 73; Mismatches 216; Indels 125; Gaps 23;

```

```

QY 4 ASKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSSTAVATAG-----QVNPIIDPMII 57
Db 1779 AGTATTASVPGTTTG-MDPGVVATTSV--TAENSSASIIAGIGGPPQQVDOQETWRT 1835
QY 58 NNFVQAPQGEFTSPNTPGGVLFDLISGLPHLNFLLHLSQMYNGVGNMRVRIWLAGNA 117
Db 1836 NFYY---NDVFTWSVADAPGSILYTVQHSPPQNNPFTAVLSQMYAGWAGMQRFVAGIG 1892
QY 118 FTAGKLIIVSCIPGFG-SHNLTIQAATLPPHVIADVRLDPIEVLDEVDVNNVLFNNDRN 176
Db 1893 VFGRLVAIVPPGIEIGPLEVRQ---FPHVVIDARSLEPVTITMPDLRENMYHTGDP 1949
QY 177 QOTMRLVCMLYTPLRTGGTGDSPVAGRVMTCPSPDNELFLVPTVEQKTRPTLPNL 233
Db 1950 GLVPTLVLSVYNNLINPFG-GSTSAIQVTETRPSEDFEFVMI-----1991
QY 237 PLSSLSNSRAPLPISGMIGSPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGTS-- 294
Db 2020 NRWNGQIVGLQ-----PVBGGF---STCNRHNLNGSTYGWSSPRFADID--- 2061
QY 350 DTFVPHLGIQANGIGSGNYIVGLSWSPSPSHSGSDVLWKIPYNGSSITEATHLAPSV 409
Db 2062 -----HRRGSAST-----PGSNATNVLOFW-YANAGSAVDNP---ISQV 2096
QY 410 YPPGFEVLVFFMSKIP--GPGAYSILPCLLPQEIYSHLASEQAPTVEGALLHYVDDTG 467
Db 2097 APQGFDP-----NSFVFPNGPG--IPA-----AGWVGFAIW---NSNSG 2131
QY 468 R-TLGBEKAYPDGFLTCVP-----NGAS-----SGPQQLPINGVEVFVSVW 507
Db 2132 APNVTTVQAYELGFATGAPGNLQPTTNTSGAQTVAKSIVAVVTGTAQNPFA-GLFVNAGCV 2190

RESULT 8
RWWRH
genome polyprotein - rabbit hemorrhagic disease virus
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C:Species: rabbit hemorrhagic disease virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 29-May-1998
C:Accession: A41039
R:Meyers, G.; Mirblich, C.; Thiel, H.J.
Virology 184, 664-676, 1991
A:Title: Rabbit hemorrhagic disease virus--molecular cloning and nucleotide sequencing o
A:Reference number: A41039; MUID:91361557; PMID:1840711
A:Accession: A41039
A:Molecule type: genomic RNA
A:Residues: 1-2344 <MEV>
A:Cross-references: GB:W67473
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: nucleotidyltransferase

Query Match 11.3%; Score 319.5; DB 1; Length 2344;
Best Local Similarity 24.2%; Pred. No. 4.7e-14;
Matches 136; Conservative 73; Mismatches 204; Indels 149; Gaps 24;

QY 4 ASKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSSTAVATAG-----VAGSSTAVATAG----- 47
Db 1779 AGTATTASVPGTTTG-MDPGVVATTSV--TAENSSASIIAGIGGPPQQVDOQETWRT 1825
QY 48 QVNPIDPWIINNVQAPQGEFTSPNTPGGVLFDLISGLPHLNFLLHLSQMYNGVGNMRV 107
Db 1826 QVDOQETWRTNFYY---NDVFTWSVADAPGSILYTVQHSPPQNNPFTAVLSQMYAGWAGM 1882
QY 108 RVRIMLAGNAFTAGKIIVSCIPGFG-SHNLTIQAATLPPHVIADVRLDPIEVLDEVDV 166
Db 1826 QVDOQETWRTNFYY---NDVFTWSVADAPGSILYTVQHSPPQNNPFTAVLSQMYAGWAGM 1882

```

```
Db 1883 QRFIVAGSGVCGRLRAVIPPGEIIGCLEVRQ---PPHVVIDARSLEPVTITMPDLR 1939
QY 167 NVLFHNDRNQOQMRVLVCMLYTLRTGGTGGDSFVVAGRVMTCPSPDFNLFPLVPP---T 223
Db 1940 PNMVHTGDPGLVPTLVLSVNNLINPFG-GSTSAIQVTVETRPSEDFEFVIRAPSSKT 1998
QY 224 VEQKTRPFTLPLNPLSSLSN-SRAPLPISGMGISPDNVQSVQFQNGRCTLDGRLVGTTPV 282
Db 1999 VDSISPAGLTTFVLTVGNGDRNRNGQIVGLQVPVGGSTC---NRHMNLNGSTYGMSSP 2055
QY 283 SLSHVAKIRGTSGTGVINLTTEL-----DGRP---PFPFEGPAPI 318
Db 2056 RCGDIDHRGASYSYSGSNATNVLQFWYANAGSAIDNPISQVAPDGFPMSPFVFNPGP--- 2112
QY 319 GPPDLG----GCDWHINMTQFGHSSQTYQYDVTPTDFVPHLSIQANGISGNGYIGVLS 374
Db 2113 GIPAGAGWVGAIWNSN-----SGAPNVTTVQAVELGPAT--GAPG 2151
QY 375 WYSPSPHSGSQVDLWKIPNYGSSITEATHLAPSVY-----PPGFEVLVTFMSK- 424
Db 2152 NLQPTTNTSQAQT-----VAKSIYAVVTGTAQNPAQ-----LFVMSG 2189
QY 425 -IPGCGAYSLPCL-LPOEYISHLASQAPTVGEAALLHYVD-----PDTGRTLGEFKAYP 477
Db 2190 IISTPNAQRITVTPQDRIVTTPTGTPAAPAVGKNTPIPFASVVRRTGDVNNATAGS----- 2244
QY 478 DGLFCLVPNGA--SSGPQQLPI 497
Db 2245 -----ANGTOYGTGSQPLPV 2259

RESULT 9
C48562
coat protein - San Miguel sea lion virus (serotype 1)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: A48562
R:Neill, J.D.
Virus Res. 24, 211-222, 1992
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea
eins.
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Accession: A48562
A:Molecule type: genomic RNA
A:Residues: 1-702 <NEI>
A:Cross-references: GB:M87481; NID:g334882; PIDN:AAA16217.1; PID:g334884
A>Note: sequence extracted from NCBI backbone (NCBIN:113564, NCBI:P:113565)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.3%; Score 261.5; DB 1; Length 702;
Best Local Similarity 23.9%; Pred. No. 1.2e-10;
Matches 137; Conservative 59; Mismatches 222; Indels 155; Gaps 26;

QY 13 DGASGAGOLVPS---VNASDPLAMPDPVAGSSAVATAGQVNPIDPWIIINFFVOAPGGEFT 69
Db 154 DPGGADIIVTERGGTVVQQPV---PAQSALTTLAAASTGKTVD-----CEWT 198

QY 70 I-----SPNTPGGLVFDLSGLPHNLPFLHLSQMYNGWGNMRVIRIMLAGNAPT 119
Db 199 TFFSYHTAVNWSSTEAQKILFSRALSPELNPLYRHISLSTWSSGIDVRFVTSGSVGF 258

QY 120 AGKIIVSCIPPGFGS-HNLTIAQATLPHVIADVRLDPIEVLDPVLEDRVNLPHNDRNQO 178
Db 259 GCKLAALIVPPGIEPVESPTMLQ---YPHVLFDARQTEPVEFTIPDIRKTLYHSD-DTD 314

QY 179 TMRVLCMLYTPL-----RTGGTGGDSFVVAGRVMTCPSPDFNLFPLVPPVTEQKTRPFTL 233
Db 315 TTRLVIMVNLINPYEQSEPSSCSITVETR-----PSSDFTFLLKPPGSLKHGSLPS 370

QY 234 PNLPLSS--LSNSRAPLPISGMGISPDNVQSVQFQNGRCTLDGRLVG-TTP-----VSL 284
```

```
Db 371 DLIPRNSRHMGNRNMWSTIDGFVQPRVFQS-----NRHFDFFDSTTTGWSPTYIPIEVTL 426
QY 285 SHV-----AKIRGTSGTGVINLTTELDTGTFPHFEGPAPIGPPDL-----GGCDW 328
Db 427 EKLDRGGQYFKYTDTEKSLVPCG-----PDGWPDTTIPMTAMTASNGYDY 471
QY 329 HI-----NMTPQ-----GHSSQT-----OYVDVDTTP 349
Db 472 TVAEYRIITNNGTHFGFYIMGNLTKVKGSDNLGETQOTSRTLFASVGNKQDQNTINPTH 531
QY 350 DTFVPHLSIQANGISGNYIGVLSWSPSPHSPSSQVDLWKIPNYGSSITEATHLA--P 407
Db 532 KITSNLVVYDANNVSAAT-AKTITHTSMHSLGYLVLDSESV---GSDSTKVRIATLUP 587
QY 408 SVYPPGFGEVLVFFMSKIPGPGAYSLPCLLPQBYISHLASQAPTVGEAALLHYVDPDG 467
Db 588 EAFYNG-GNFPVFFTNKI-----QIGHFDRAHTKCFNSQVLM-----TS 625
QY 468 RTLGE--FKAYDPDGLTCLVPNGASSGPQOLPIN 498
Db 626 OKLAENHVTLPDPSLLVYRITDAASSWFDLGIN 658

RESULT 10
C48562
coat protein - San Miguel sea lion virus (serotype 4)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: C48562
R:Neill, J.D.
Virus Res. 24, 211-222, 1992
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea
eins.
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Accession: C48562
A:Molecule type: genomic RNA
A:Residues: 1-703 <NEI>
A:Cross-references: GB:M87482; NID:g334886; PIDN:AAA16220.1; PID:g334888
A>Note: sequence extracted from NCBI backbone (NCBI:P:113567)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 8.9%; Score 251; DB 1; Length 703;
Best Local Similarity 25.1%; Pred. No. 6.4e-10;
Matches 116; Conservative 67; Mismatches 188; Indels 92; Gaps 23;

QY 13 DGASGA-----GOLVPEVNASDDPLAMPDPVAGSSAVATAGQVNPIDPWI-----I 57
Db 154 DPGSABEIVTEEGTVVQQPAPAPATLATLATATASTGKSVEQE-----WMTFFSVHTSI 207

QY 58 NNFOAPQGEFTISPNTPGVLPDLSGLPHNLPFLHLSQMYNGWGNMRVIRIMLAGNA 117
Db 208 NMWSTVESQCK-----ILYSQALNPSINPLYDHIKLYSTWSSGIDVRFVTSGS 256

QY 118 FTAGKIIVSCIPPGFGS-HNLTIAQATLPHVIADVRLDPIEVLDPVLEDRVNLPHNDRN 176
Db 257 VFGGKLAALLVPPGVEPIESVMLQ---YPHVLFDARQTEPVEFTIPDIRKTLFHSMDT 313

QY 177 QOTMRLVCMLYTLPRLTGGTGGDSFVVAGRVMTCPSPDFNLFPLVPP--TVEQKTRPFTLP 234
Db 314 -DTTKLVINPYE-----NGVENKTTCSITVETRPASADFTFALLKPPGSLKHGSLPSDL- 366

QY 235 NLPGLSL--SNSRAPLPISGMGISPDNVQSVQFQNGRCTLDGRLVG-TTPVSLSHVAKIR 291
Db 367 -IPRNSAHMGNRNMWSTISGFSVQPRVFQS-----NRHFDFFDSTTTGWSPTYIPIEIKI 421

QY 292 GT--SNGTIVINLTDELDTGTFPHFEGPAPIGPPDLGGCDWHINMTQFGHSSQTYQYDVTTP 349
Db 422 GKVGNNKWFHVIDTD-----KALVGPIDGWPDDTIPD-ETKATNGNFSYGESYRAGST- 475
```


QY 350 DTFVPHLSIQANGICSGNYI-GVLSWVSPSPSH-----PSGSQVDLWKIPNYGSSI 399
 Db 476 -TIKENENSTHF----KGTTCIGTSTVEIPENDSQQIKTEAKKSKQTMVYVVTADFKDTI 530
 QY 400 TEATH-LAPSVYPPGEGVLVVF-----MSKIPGGAYSL 433
 Db 531 VKPQHKISQ-----KLWVYFDGPEKDLTMSATLSPLGVTL 566

RESULT 11
 JQ2354
 coat protein - feline calicivirus (strain NADC)
 C:Species: feline calicivirus
 C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
 C:Accession: JQ2354
 R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.
 J. Gen. Virol. 74, 2519-2524, 1993
 A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable
 A:Reference number: JQ2354; MUID:94065683; PMID:7504075
 A:Accession: JQ2354
 A:Molecule type: mRNA
 A:Residues: 1-668 <SEA>
 A:Cross-references: GB:I09718; NID:G305104; PIDN:AAA16485.1; PID:G305105
 C:Superfamily: feline calicivirus coat protein
 C:Keywords: capsid protein; coat protein

Query Match 8.6%; Score 242.5; DB 2; Length 668;
 Best Local Similarity 32.2%; Pred. No. 2.4e-09;
 Matches 64; Conservative 27; Mismatches 83; Indels 25; Gaps 6;
 QY 34 DPVAGSSTAVATAGQVNPDPWLNINNVQAPQGEFTI-----SPNTPGGVLDLSLGP 87
 Db 145 EPSAQMSSTAADMATGSKVDSEW-----EAFSFTSVNWNSTSETQKILFKQSLGP 195
 QY 88 HLNPFLLHLSQMYNGVGNMVRIMLAGNAFTAGKIIVSCIPPGFSGHNLITIAQATLFFH 147
 Db 196 LLNPLYLHLSKLYVAMSGSVEVRFISGSGVFGKLAALIVPPGVDPVQST--SMLQYPH 253
 QY 148 VIADVRTLDPIEVLDRVNLPHNNDRNQOTMRLVCMLY-----TPLRTGGGTGDSFVA 203
 Db 254 VLPDARQVDPVIFSPIDRLSTLYHLMF-DTDTTSLVIMVYNDLINPYANDSNSSGCIVT- 311

QY 204 GRVMTCPSPDPNFLFLVPP 222
 Db 312 --VETKPGDPKFHLLKPP 328
 RESULT 12
 VCWNEC
 coat protein - feline calicivirus (strain Japanese F4)
 N:Alternate names: capsid protein
 C:Species: feline calicivirus
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
 C:Accession: B40481
 R:Tohya, Y.; Taniguchi, Y.; Takahashi, E.; Utagawa, E.; Takeda, N.; Miyamura, K.; Yamaza
 Virology 183, 810-814, 1991
 A:Title: Sequence analysis of the 3'-end of feline calicivirus genome.
 A:Reference number: A40481; MUID:91306470; PMID:1853578
 A:Accession: B40481
 A:Molecule type: genomic RNA
 A:Residues: 1-668 <TOH>
 A:Cross-references: GB:D90357; NID:G221264; PIDN:BAAL14371.1; PID:G221266
 C:Superfamily: feline calicivirus coat protein
 C:Keywords: capsid protein; coat protein; glycoprotein
 C:Keywords: capsid protein; coat protein; glycoprotein (Asn) (covalent) #status predicted
 F:177,301,304,399,459,465/Binding site: carbohydrate (Asn)

Query Match 8.6%; Score 241.5; DB 1; Length 668;
 Best Local Similarity 31.7%; Pred. No. 2.8e-09;
 Matches 63; Conservative 28; Mismatches 83; Indels 25; Gaps 6;
 QY 34 DPVAGSSTAVATAGQVNPDPWLNINNVQAPQGEFTI-----SPNTPGGVLDLSLGP 87
 Db 145 EPSAQMSSTAADMATGSKVDSEW-----EAFSFTSVNWNSTSETQKILFKQSLGP 195

QY 88 HLNPFLLHLSQMYNGVGNMVRIMLAGNAFTAGKIIVSCIPPGFSGHNLITIAQATLFFH 147
 Db 196 LLNPLYLHLSKLYVAMSGSVEVRFISGSGVFGKLAALIVPPGVDPVQST--SMLQYPH 253
 QY 148 VIADVRTLDPIEVLDRVNLPHNNDRNQOTMRLVCMLY-----TPLRTGGGTGDSFVA 203
 Db 254 VLPDARQVDPVIFSPIDRLSTLYHLMF-DTDTTSLVIMVYNDLINPYANDSNSSGCIVT- 311
 QY 204 GRVMTCPSPDPNFLFLVPP 222
 Db 312 --VETKPGDPKFHLLKPP 328

RESULT 13
 VCWNEF
 coat protein - feline calicivirus (strain CFI/68 FIV)
 N:Alternate names: capsid protein
 C:Species: feline calicivirus
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Aug-1999
 C:Accession: A40507; B40507; T09246
 R:Neill, J.D.; Reardon, I.M.; Heinrichson, R.L.
 J. Virol. 65, 5440-5447, 1991
 A:Title: Nucleotide sequence and expression of the capsid protein gene of feline caliciv
 A:Reference number: A40507; MUID:91374597; PMID:1716692
 A:Accession: A40507
 A:Molecule type: genomic RNA
 A:Residues: 1-668 <NEI>
 A:Cross-references: GB:W32819; NID:G323874; PIDN:AAA42925.1; PID:G323875
 A:Accession: B40507
 A:Molecule type: protein
 A:Residues: 373-379; 403-419; 481-489; 560-566 <NE2>
 R:Neill, J.D.
 submitted to the EMBL Data Library, April 1998
 A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.
 A:Reference number: Z16626
 A:Accession: T09246
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: genomic RNA
 A:Residues: 1-668 <NE3>
 A:Cross-references: EMBL:U13992; NID:G3056875; PIDN:AAC13993.1; PID:G537256
 A:Experimental source: strain CFI/68 FIV
 C:Superfamily: feline calicivirus coat protein
 C:Keywords: capsid protein; coat protein; glycoprotein
 F:177,301,304,439,459,465/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:177,301,304,439,459,465/Binding site: carbohydrate (Asn)

Query Match 8.5%; Score 240.5; DB 1; Length 668;
 Best Local Similarity 29.9%; Pred. No. 3.3e-09;
 Matches 64; Conservative 29; Mismatches 86; Indels 35; Gaps 7;
 QY 19 GOLVPEVNASDPLAMDVPVAGSSTAVATAGQVNPDPWLNINNVQAPQGEFTI-----SP 72
 Db 140 GGIVAEPAQMSSTAADMATGSKVD-----SEW-----EAFSFTSVNWNST 180
 QY 73 NNTPGGVLDLSLGPLHLPFLHLSQMYNGVGNMVRIMLAGNAFTAGKIIVSCIPPGF 132
 Db 181 SETQKILFKQSLGPLHLPFLHLSQMYNGVGNMVRIMLAGNAFTAGKIIVSCIPPGF 240
 QY 133 GSHNLITIAQATLFFHVIADVRTLDPIEVLDRVNLPHNNDRNQOTMRLVCMLY-----T 188
 Db 241 DPVQST--SMLQYPHVLPDARQVDPVIFSPIDRLSTLYHLMF-DTDTTSLVIMVYNDLIN 297
 QY 189 PLRTGGGTGDSFVAGRVMTCPSPDPNFLFLVPP 222
 Db 298 PYANDSNSSGCIVT---VETKPGDPKFHLLKPP 328

RESULT 14
 VCWNEF9
 coat protein - feline calicivirus (strain F9)
 N:Alternate names: capsid protein
 C:Species: feline calicivirus
 C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 28-Jul-2000

Query Match	8.1%	Score 229.5;	DB 2;	Length 668;
Best Local Similarity	29.4%;	Pred. No. 2e-08;		
Matches 63;	Conservative	28;	Mismatches 88;	Indels 35; Gaps 7;

Search completed: June 1, 2004, 13:55:18
Job time : 11.2265 secs

THIS PAGE BLANK (USPTO)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2033.5	72.1	546	1	COAT_SOUV3	Q04542 southampton
2	1150.5	40.8	539	1	COAT_LORDV	P54635 lordsdale v
3	320.5	11.4	2344	1	POLN_RBDV	P27410 rabbit hemo
4	261.5	9.3	702	1	COAT_SNSV1	P36284 san miguel
5	252	8.9	2208	1	POLN_MANCV	Q69014 manchester
6	251	8.9	703	1	COAT_SNSV4	P36285 san miguel
7	241.5	8.6	668	1	COAT_FCVF4	P27405 feline cali
8	240.5	8.5	668	1	COAT_FCVF6	P27404 feline cali
9	237	8.4	671	1	COAT_FCVF9	P27406 feline cali
10	149	5.3	2205	1	POLG_POL2W	P23069 p genome po
11	147	5.2	2207	1	POLG_POL2L	P06210 p genome po
12	142	5.0	2194	1	POLG_HE701	P32537 h genome po
13	140.5	5.0	2206	1	POLG_POL32	P06209 poliovirus
14	138	4.9	2153	1	POLG_HRV16	Q83122 h genome po
15	136.5	4.8	2185	1	POLG_SVDVU	P13900 s genome po
16	134.5	4.8	2164	1	POLG_HRV89	P07210 h genome po
17	132	4.7	2193	1	POLG_HE71B	Q68479 h genome po
18	131	4.6	2193	1	POLG_HE71B	Q68478 h genome po
19	130.5	4.6	2432	1	Y43R_IRV6	P18305 chilo iride
20	128.5	4.6	2185	1	POLG_SVDVH	P16604 s genome po
21	128	4.5	855	1	POLG_HRV3	Q82081 human rhino
22	127.5	4.5	2175	1	POLG_BOVEV	P12915 b genome po
23	126.5	4.5	2196	1	POLG_EC05N	Q9Y1J1 e genome po
24	126	4.5	2185	1	POLG_CXB5P	Q30353 c genome po
25	126	4.5	2206	1	POLG_CXA21	P22055 c genome po
26	126	4.5	2206	1	POLG_POL1M	P03299 p genome po
27	124.5	4.4	1324	1	SALI_HUMAN	Q9N5C2 homo sapien
28	123.5	4.4	862	1	POLG_EC16H	Q66790 echovirus 1
29	123	4.4	1200	1	HYAL_SF8PU	O76536 strongyloce
30	122	4.3	2206	1	POLH_POL3L	P03302 poliovirus
31	122	4.3	2208	1	POLH_POL1M	P03300 p genome po
32	120.5	4.3	1183	1	DRPL_RAT	P54258 rattus norv
33	120	4.3	1477	1	VL2_Hpv42	P27235 human papil

Matches 249; Conservative 82; Mismatches 182; Indels 39; Gaps 14;

QY 1 MMASKDATTSSVDGASGAGQVPEVNAQDPLAMDPAVAGSTAVATAGVNPIDPWIINNF 60
 DB 1 MKMASNDANPS-DGS--AANLVPEVN-NEVMALEFPVGAATAAPVAGQONVIDPWIRNF 56
 QY 61 VQAPQCEFTISPNNTPGGVLFDLSLPHLNPFLHLSOMYNGVGMVRIMLAGNATA 120
 DB 57 VQAPGGEFTVSPNAPGEILWSAPGLDPLNPLSHLSRMVYAGGFEVQVILAGNAFTA 116
 QY 121 GKIIVSCIPPGFGSHNLTIAQATLPHVITADYRLDPIEVLPLEDVRNVLPHNDRNQTM 180
 DB 117 GKVIFAAPNPFTPEGLSPSQVTMPFHIIIVDVRQLEPILIPDVRNNFYHYNQANDSTL 176
 QY 181 RLVMCLYTLRTGGTGGDSFVWAGRVMTCPSPDFENFLVPPPTVEQKTRPFLPNLPLSS 240
 DB 177 KLFAMLYTPURANNAGDDVFTSCRVLTPSPDFDFELVPPPTVESRTKPTVPVLTVBE 236
 QY 241 LNSRAPLPTISGMGIPDNVQSVQFQNGRCTLDRGLVGTTPVSLSHVAKIRGT-----S 294
 DB 237 MNSRPPLEKLYTGPSSAFVVPQNGRCTLDRGLVGTTPVSLSHVAKIRGT-----S 296
 QY 295 NGTVINLTDLGTFPHFEG-PAPIGPDLGGCDHINMTQF----CHSSQTOQYDVTTP 349
 DB 297 HDYTMNLASQNSNYDPTIEIPAPLGTPTDFVGIQGL-LTQITRADGSTRAHKATVSTGS 355
 QY 350 DTFVPHLGSIQ-----ANGIGSGNY-----IGVLSWSPSPHSPSGQVDLWKIPNYGSSI 399
 DB 356 VHTFPLKGSVQFTDNNDFQAGNKTFTVGVQIQ--DGDHQNQEP--QMSLPNYSGR 411
 QY 400 TEATHLAPSVPPGFGVILVFFMSKIPGAY---SLPCLLPQSYISHLASEQAPTVEA 456
 DB 412 GHNVLAPAVAPTPEGQLLFFRSTWPGCGYGNMNLDCLLPQBWVLHFQEAAPQSDV 471
 QY 457 ALLHYVDPDGTIGEFKAYPDGFLTCVPGNGASSGPOQ--LPINGVFVFWVSWRFRYQLK 514
 DB 472 ALLRFVAPDTRGVLFCEKLLHKSGLVTV---AHTGPDVLDLPPNGYFRFSDWVNFYTLA 527
 QY 515 FVGTASSARGRL 526
 DB 528 PMNGTGRRL 539

RESULT 3
 ID POLN RHDV STANDARD; PRT; 2344 AA.
 AC P27410;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Non-structural polyprotein [Contains: RNA-directed RNA polymerase
 DE (EC 2.7.7.48); Thiol protease P3C (EC 3.4.22.-); Helicase (2C like
 DE protein); Coat protein].
 DE Rabbit hemorrhagic disease virus (RHDV).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Lagovirus.
 OX NCBI_TaxID=11976;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91361557; PubMed=1840711;
 RA Meyers G., Wirblich C., Thiel H.-J.;
 RT "Rabbit hemorrhagic disease virus -- molecular cloning and nucleotide
 RT sequencing of a calicivirus genome.";
 RL Virology 184:664-676(1991).
 CC -/- FUNCTION: THE P3C CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE
 CC FOR CLEAVAGE AT CERTAIN Q/G SITES IN THE POLYPROTEIN.
 CC -/- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC [RNA] (N).
 CC -/- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

QY 181 RLVMCLYTLRTGGTGT--DSFVWAGRVMTCPSPDFENFLVPPPTVEQKTRPFLPNLPL 238
 DB 180 RLVMCLYTLRTGGSGNSDSFVWAGRVLTAPSSDFSLFLVPPPTIEKTRAFVPPNIPL 239
 QY 239 SLSNSRAPLPTISGMGIPDNVQSVQFQNGRCTLDRGLVGTTPVSLSHVAKIRGTSN--G 296
 DB 240 QTLNSNRPSLIQGMILSPDASQVVFQNGRCLIDQLLGTTPATSGQLFRVGRKINGQA 299
 QY 297 TVINLTDLGTFPHFEGPAPIGPDLGGCDHINMTQFHS-----QTQVDVDTTPTD 351
 DB 300 RTNLNTEVDGKPFMAFDSAPVGFDFGKCDHMRKISKTPNNTGSDPMRSVSVQTNVQG 359
 QY 352 FVPHLGSIQANGI--GSGNYIGVLSWSPSPHSPSGQVDLWKIPNYGSSITEATHLAPS 408
 DB 360 FVPHLGSIQDFVENHPTGDYGTIEWISQPTPTGTDINLWEIPDYGSSLSQAANLAPP 419
 QY 409 VYPPGFGVILVFPMSKIPGAYS-----LPCLLPQSYISHLASEQAPTVEAALHYVDP 464
 DB 420 VFPFGFEALVFPVSAFPGPNRSAPNDVFCLLPQBYIIFVSEQAPTMDALHYVDP 479
 QY 465 DTGRTIGEFKAYPDGFLTCVPGNGASSGPOQLPINGVFWVSWRFRYQLKPVGTASSARG 524
 DB 480 DTRNRLGEEFLYEGGILTCVPGNGAGAGPQQLPLNGVLFVSWRFRYQLKPVGTASTARG 539
 QY 525 RLGLRR 530
 DB 540 RLGVRR 545

RESULT 2
 ID COAT LORDV STANDARD; PRT; 539 AA.
 AC P54635;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Coat protein (Capsid protein).
 DE Lordsdale virus (Human enteric calicivirus).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 OX NCBI_TaxID=82658;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96005060; PubMed=7561776;
 RA Dingle K.E., Lambden P.R., Caul E.O., Clarke I.N.;
 RT "Human enteric Caliciviridae: the complete genome sequence and
 RT expression of virus-like particles from a genetic group II small
 RT round structured virus.";
 RL J. Gen. Virol. 76:2349-2355(1995).
 CC -/- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X86557; CAA60255.1;
 DR InterPro; IPR004005; Calici.coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 539 AA; 58775 MW; 5E5C63E7F2C5FD21 CRC64;
 Query Match 40.8%; Score 1150.5; DB 1; Length 539;
 Best Local Similarity 45.1%; Pred. No. 2.6e-75;

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC EMBL; M67473; AAA47285.1; -
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR000317; Peptidase_C24.
 DR InterPro; IPR000605; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_P5vir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 DR Pfam; PF03510; Peptidase_C24; 1.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00916; 2CENDOPTASE.
 DR PRINTS; PR00918; CALICIVIRUSNS.
 DR SMART; SM00382; AAA; 1.
 KW Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;
 KW Thiol protease; Helicase; ATP-binding; Coat protein; 3D-structure.
 FT CHAIN 7508 7707
 FT CHAIN 1109 1251
 FT CHAIN 71497 71625
 FT CHAIN ? 2344
 FT NP_BIND 522 529
 FT ACT_SITE 1212 1212
 FT ACT_SITE 1227 1227
 FT ACT_SITE 2344 2344
 SQ SEQUENCE 2344 AA; 257066 MW; 1454C248F81E9212.CRC64;

Query Match 11.4%; Score 320.5; DB 1; Length 2344;

Best Local Similarity 24.2%; Pred. No. 8.4e-15;

Matches 136; Conservative 74; Mismatches 203; Indels 149; Gaps 24;

QY 4 ASKDATSSVDGASGAGQLVPEVNASDPLAMDPR-----VAGSSTAVATAG----- 47
 DB 1779 ACTATTASVPGTTDGT-----MDPGVAVTTSVITAENSSASIAATGAGPPQ 1825
 QY 48 QVNPDPWIIINNFVQAPQCEFTISPNTPGGVLFDSLGLPHLPFLHLSQMYNGWVGNM 107
 DB 1826 QVQDQETWTFNY---NDVFTWSVADAPGSLTYVQHSPPNPPTAVLSQMYAGWGM 1882
 QY 108 RVRIMLAGNAFTAGKIIIVSCIPPGFG-SHNLTAQATLPPHVIADVRLDPIEVLPLEV 166
 DB 1883 QPRFIVAGSGVGGRLVRAVIPPGEIGGLEVRQ---PPHVVIDARSLEPVTITMPDLR 1939
 QY 167 NVLFHNDRNQMTMLVCMYLTPLRTGGTGDGDFVAGRVMTCPSPDFNLFVLP---T 223
 DB 1940 PNYHTPGDGLVPTLVLSVNNLNPFG-GSTSAIQVTVTRPSEDEFVIMRAPSSKT 1998
 QY 224 VEQKTRPFTLPNPLSSLSN-SRAPLPSGMGIPDNVQSVQFQNGRCTLDGLRGVTPV 282
 DB 1999 VDSISGAGLLTTPVTGNDNRNNGQIVGLQVPGGFGSTC---NRHNLNGSTVGNSSP 2055
 QY 283 SLISHVAKIRGTSNGTVINLTTEL-----DGTPT---FHPFEGPAPI 318
 DB 2056 REGDIDHRGASYSGSNATNLVQFWYANAGSAIDNPISQVAPDGFDMSPFVFNPG--- 2112
 QY 319 GPPDLG-----GCDWHINMTQFCHSSQTQYDVDTTDPFVPHLGSTQANGIGSGNVIG 374
 DB 2113 GIPAGAGWVGFAIWSN-----SCAPNTVTVQAVELGPAT---GAPG 2151
 QY 375 WYSPSPHSGSVDLWKIPNYGSSITEATHLAPSVY-----PPGGEVLVLFMSK- 424
 DB 2152 NLQPTINTSGAQT-----VAKSIYAVVTGAQNAPG-----LFFMASG 2189

QY 425 -IPPGGAYSLPCL-LPQBYISHLASEQAPTYCEAALLHYVD-----PDTGRTLGEFKAYP 477
 DB 2190 IISTENASAITVTPQPDRIIVTPGTPAAAPVGKNTPIMFASVVRRTGVDVNATAGS----- 2244

QY 478 DGFLTCVPNGA--SSGPQOLPI 497
 DB 2245 -----ANGTQYGTGSQPLPV 2259

RESULT 4

COAT_SMSV1
 ID_COAT_SMSV1 STANDARD; PRT; 702 AA.
 AC P36284;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Coat protein (Capsid protein).
 OS San Miguel sea lion virus (serotype 1) (SMSV 1).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OX NCBI_TaxID=36406;
 RN [1]
 RX SEQUENCE FROM N.A.
 RP MEDLINE=92410750; PubMed=1529644;
 RA Neill J.D.;
 RT "Nucleotide sequence of the capsid protein gene of two serotypes of
 RT San Miguel sea lion virus: identification of conserved and non-
 RT conserved amino acid sequences among calicivirus capsid proteins.";
 RL Virus Res. 24:211-222(1992).
 CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M87481; AAA16217.1; -
 DR PIR; A48562; A48562.
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 481 481 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 493 493 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 545 545 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 702 AA; 77850 MW; E6E5A58523DEE3D7.CRC64;

Query Match

Best Local Similarity 9.3%; Score 261.5; DB 1; Length 702;

Matches 137; Conservative 59; Mismatches 222; Indels 155; Gaps 26;

QY 13 DGASGAGOLVPE---VNASDPLAMDPLVAGSSTAVATAGQVNPDPWIIINNFVQAPQGEFT 69
 DB 154 DPGGADIVTEEGTIVVQQFV---PAQSALTTLAAASTGKTVD-----CEWT 198
 QY 70 I-----SPNTPGGVLFDSLGLPHLPFLHLSQMYNGWVGNVRIMLAGNAFT 119
 DB 199 TFFSYHTAVNWSSTEAQKILFSRALSPELNPYLRIHISLSYTWSSGGIDVRFVTSGSGVF 258
 QY 120 AGKIIIVSCIPPGFGS-HNLTAQATLPPHVIADVRLDPIEVLPLEVDRNVLPHNDRNQ 178
 DB 259 GCKLAALVPPGIEPVESPTMLQ---YHVLFDARQTEPVTFTIDIKTLVHSD-DTD 314
 QY 179 TMRVLCMLYTEL-----RTGGTGDGDFVAGRVMTCPSPDFNLFVLPPTVEQKTRPFTL 233
 DB 315 TTRLVIMVYNELINPYEQSEPKSSCSITVETR---PSSDFTFSLKPKPGLLKHGSI 370
 QY 234 PNLPLSS--LNSRAPLPIISGMGISPDNVQSVQFQNGRCTLDGLRGV-TTP-----VSL 284

Db 371 DLIPRNSRHWGNRHWSTIDGFFVQVRVFS-----NRHFDSDSTTCWSTPYIPIEVTIL 426
 QY 285 SHV-----AKIRGTSNGTVINLTDLGTPHPFEGPAPIGFPPDL-----GGCDW 328
 Db 427 EKLDRGGQVFKVTDTEKSLVPL-----PDGWPDTTPTANTANGNYDY 471
 QY 329 HI-----NMQF-----GHSSOT-----QYDVTTP 349
 Db 472 TVAEVRIITNNGTHFGFYMGHLTTKVKGSNDLGETQTSRTLPASVGNVQDQNTINPTH 531
 QY 350 DTFVPHLGSIOANGIGSGNYIGVLKVSPPSPHSPGSQVDLWKIPNYSIITEATHLA--P 407
 Db 532 KITSNSLVVYDANNVSAAT-AKTTTWHSTMSHLGYLVVDESVP-----GSDSTKVRIATLP 587
 QY 408 SVYPPGFGVLFVFFMSKIPGPGNYSLPCLLPQEXYLSHLASEQAPTVGEAALLHYVDPTG 467
 Db 588 EAFNG-GNFPVFETTKI-----QGHFDRANTKFNCSQVLM-----TS 625
 QY 468 RTLGE--FRAYPDGFLTCVPNGASSGPQQLPIN 498
 Db 626 QKLAENHYTLPPDSLIVYRITDAASSWFDLGIN 658

RESULT 5
 ID POLN MANCV STANDARD; PRT; 2208 AA.
 AC Q69014; 1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polyprotein [Contains: RNA-directed RNA polymerase
 (EC 2.7.7.48); Thiol protease 3C (EC 3.4.22.-); Helicase (2C like
 protein); Coat protein].
 DE Manchester virus (human enteric calicivirus).
 OS Manchestervirus (human enteric calicivirus).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Sapovirus.
 NCBI_TaxID=82659;
 RX SEQUENCE FROM N.A.
 RX MEDLINE=95390791; PubMed=7661689;
 RA Liu B.L., Clarke I.N., Caul E.O., Lambden P.R.;
 RT "Human enteric caliciviruses have a unique genome structure and are
 distinct from the Norwalk-like viruses.";
 RL Arch. Virol. 140:1345-1356(1995).
 CC -!- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).
 CC -!- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR
 THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY
 SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 {RNA}(N).
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; X86560; CAA60262.1; --
 CC MEROPS; C24.UPW; --
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR00317; Peptidase_C24.
 DR InterPro; IPR00605; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.

DR Pfam; PF03510; Peptidase_C24; 1.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PRO0916; 2CENDOPTASE.
 DR PRINTS; PRO0918; CALICIVIRUSNS.
 KW Polyprotein; transferase; RNA-directed RNA polymerase; Hydrolase;
 KW Thiol protease; Helicase; ATP-binding; Coat protein.
 KW CHAIN ? ? HELICASE (P2C-LIKE).
 FT CHAIN ? ? PROTEASE (P3C).
 FT CHAIN ? ? RNA-DIRECTED RNA POLYMERASE.
 FT CHAIN ? ? COAT PROTEIN.
 FT NP_BIND 408 415 ATP (POTENTIAL).
 FT ACT_SITE 1097 1097 PROTEASE (BY SIMILARITY).
 FT ACT_SITE 1112 1112 PROTEASE (BY SIMILARITY).
 FT ACT_SITE 2208 AA; 242736 MW; 3E299D5BA20E45DC CRC64;
 SQ SEQUENCE 2208 AA; 242736 MW; 3E299D5BA20E45DC CRC64;
 Query Match 8.9%; Score 252; DB 1; Length 2208;
 Best Local Similarity 24.3%; Pred. No. 6.5e-10; Indels 58; Gaps 11;
 Matches 99; Conservative 57; Mismatches 193;
 QY 17 GAGQLVPEVNASDPLAMDVPVAGSSTAVATAGQVNPIDP-----WIINNFVQ 62
 Db 1650 GNGSNPEPKOSNPMVVDPPGTTGTTSHVWVNPQNGAAQRLAELAVATGAIOSNVE 1709
 QY 63 APQGEFTI-----SPNTPGGVLFDLGLPHLNPFLHLSQMYNGWGNMVRIMLAGN 116
 Db 1710 AIRNCPAVERTEAFANDRMPTGTFLGSLHPNINPYTSHLSGMWAGMGSGFEVRLSIS 1769
 QY 117 AFTAGKLIIVSCIRPGFSGHNLTAQATLPHVIAVRTLDPIEVPLEDVRLVFNHNDNRN 176
 Db 1770 GVFAIRIIASVIPPVVDPS--SIRDPGVLPFAVDARITEPVSMIPDVRADVHYRMDGA 1827
 QY 177 QQTMLRLVCLMYTLPLRTGGTGGDSFVWAGRYMTCPSPDFNLFVLP--TVEQKTRFTLP 234
 Db 1828 EPTCSLGFYVYQPLLNPFTTAVSTCWVSVEKPGGDFDCLLRPPQQQMGNGVSEGLL 1887
 QY 235 NLPLSSLSNRAPLPISGMGISPDNVQ--SVQFQNGRCTLGRLVGTTPVLSLHVA--KIRG 292
 Db 1888 PRRLGYSRGNRVGLVVGMLVAEHKQVNRHNSNSVTFGWSVAPVNPMAAEIVTNOAHS 1947
 QY 293 TSNQTVINLTDLGTPHPFEGPAPIGFPP-----LGG-----CDWHINMT 333
 Db 1948 TSHRWLSIGQKNGKPLP---GIPNHPDSCASTVVGAMDTSLGSRPSTGVCGPALISFQ 2004
 QY 334 QGHSSQYQYDVDTTPTDTPFVPHLSIQANGISGNYIGVLVSWVSPPS 380
 Db 2005 NNG-----DVYENDTFSVMFATY-----DPLTSGTGVALTNSINPAS 2041
 RESULT 6
 ID COAT SMSV4 STANDARD; PRT; 703 AA.
 AC P36285;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Coat protein (Capsid protein).
 OS San Miguel sea lion virus (serotype 4) (SMSV 4).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 NCBI_TaxID=36407;
 RX SEQUENCE FROM N.A.
 RX MEDLINE=92410750; PubMed=1529644;
 RA Neill J.D.;
 RT "Nucleotide sequence of the capsid protein gene of two serotypes of
 San Miguel sea lion virus: identification of conserved and non-
 conserved amino acid sequences among calicivirus capsid proteins.";
 RL Virus Res. 24:211-222(1992).
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; M87482; AA16220.1; --
 CC DR PIR; C48562; C48562; Calici_coat.
 CC DR InterPro; IPR004005; Calici_coat.
 CC DR Pfam; PF00915; Calici_coat; 1.
 CC KW Coat protein; Glycoprotein.
 CC FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC FT CARBOHYD 482 482 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC SQ SEQUENCE 703 AA; 77721 MW; C5DAD8223B261073 CRC64;

Query Match 8.9%; Score 251; DB 1; Length 703;
 Best Local Similarity 25.1%; Pred. No. 1.7e-10;
 Matches 116; Conservative 67; Mismatches 188; Indels 92; Gaps 23;

QY 13 DGASGA-----GQLVPEVNASDPLAMPVAGSSSTAVATAGQVNPIDPW-----I 57
 DB 154 DPGSAEIVTEEGTVQOQPAPATLATATASTGKSVEQE-----WMTFFSYHTSI 207
 QY 58 NNFVQAQPGQFTSPNNTGGVFLDLSGLPHLPNPFLLHLSOMYNGVGNMVRIMLAGNA 117
 DB 208 NWSVESQK-----ILYSQALNPSINPDLHIAKLYSTWSGGIDVRFVSGSG 256
 QY 118 FTAGKIIVCIPPGFS--NLTIAQATLPHPHVIADVRLDPIEVLPLEVRLNHNDRN 176
 DB 257 VFGKLAALLVPPGVEPIESVMLQ---YPHVLFARQTEPVIPTIPDIRKTLFHSMDST 313
 QY 177 QOTRLVCLMYPLRTGGGTGSFVAGVMTCPSPDFNLFVLP--TVQKTRPTLP 234
 DB 314 -DTTKLVINPYE-----NGVENKTCITVETRPSADFTFALLKPGSLIKGSPSLL- 366
 QY 235 NPLSSL--SNSRAPLPISGMGISPDNVQSFQNGRCRTLDGLRGV--TTPVSLSHVAKIR 291
 DB 367 -IPRNSAHMGNWSTISGFSVQPRVFS-----NRHFDFTSTTGWSTPYVPIEIKIQ 421
 QY 292 GT--SNGTVINTELDTGTFHPPEGPAPIGFDLGGCDWHNNQTHGHSQTYQVDVTPP 349
 DB 422 GKVSNNKWFHVIDT---KALVPGIPDGPDPDTPID--ETKATNGNFSYGESYRAGST- 475
 QY 350 DTFVPHLSIQANGISGNYI--GVLSWVSPSH-----PSGSQVLDLWKPVNGSSI 399
 DB 476 -TIKPNENSTHF-----KGTIYICGLSTVBEIPENDEQOIKTEAPKKSQTMVYVVTADPKDTI 530
 QY 400 TEATH-LAPSVYPPGFGVLPF-----MSKIPGPGAYSL 433
 DB 531 VKPQHKISQ-----KLIVYFDGPEKDLTWSATLSPLCYTL 566

RESULT 7

COAT_FCVF4
 ID COAT_FCVF4 STANDARD; PRT; 568 AA.
 AC F27405;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Coat protein (Capsid protein).
 OS Feline calicivirus (strain Japanese F4) (FCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OX NCBI_TaxID=11980;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91306470; PubMed=1853578;
 RA Tohya Y., Taniguchi Y., Takahashi E., Utagawa E., Takeda N.,
 RA Miyamura K., Yamazaki S., Mikami T.;

RT "Sequence analysis of the 3'-end of feline calicivirus genome.";
 RL Virology 183:810-814(1991).
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALCIVIRUSES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; D90357; BAAL4371.1; --
 CC DR PIR; B40481; VCMWPC.
 CC DR InterPro; IPR004005; Calici_coat.
 CC DR InterPro; IPR008975; Viral_Cap_coat.
 CC DR Pfam; PF00915; Calici_coat; 1.
 CC KW Coat protein; Glycoprotein.
 CC FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC SQ SEQUENCE 668 AA; 73589 MW; 85BDBC85804E503 CRC64;

Query Match 8.6%; Score 241.5; DB 1; Length 668;
 Best Local Similarity 31.7%; Pred. No. 7.8e-10;
 Matches 63; Conservative 28; Mismatches 83; Indels 25; Gaps 6;

QY 34 DPVAGSSSTAVATAGQVNPIDPWIIINNVQAQPGQFTI-----SPNTPGGVLPDLSLGP 87
 DB 145 EPSAQMAADWASGKSDSEW-----EAFSFTSVNMWSTSETQKILFKQSLGP 195
 QY 88 HLNPELLHLSOMYNGVGNMVRIMLAGNAPTAKIIVCIPPGFSGSHNLTAQATLPH 147
 DB 196 LNPVLEHLKUYVWAGSIEVRFSIGVGKLAALVFPVGPVQST--SMLOIYH 253
 QY 148 VIADVRLDPIEVLPLEVRLNHNDRNQOTRLVCLMY-----TPLRTGGGTGSFVVA 203
 DB 254 VLFDARQVEPVIPTIPDLRSTLYHWS-DTDTTSLVIMVYNDLINPYANDSSGCIIVT- 311
 QY 204 GRVMTCPSPDFNLFVLP 222
 DB 312 --VETKPGDPKFKHLLKPP 328

RESULT 8

COAT_FCVG6
 ID COAT_FCVG6 STANDARD; PRT; 668 AA.
 AC P27404;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Coat protein (Capsid protein).
 OS Feline calicivirus (strain CFI/68 FIV) (FCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OX NCBI_TaxID=11979;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=91374597; PubMed=1716692;
 RA Neill J.D., Reardon I.M., Heinrichson R.L.;
 RT "Nucleotide sequence and expression of the capsid protein gene of
 RT feline calicivirus";
 RL J. Virol. 65:5440-5447(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Neill J.D.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- PTM: The N-terminus is blocked.
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALCIVIRUSES.
 CC -----

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M12197; AAA46912.1; -;
 DR PIR; A29507; GNNVSP.
 DR HSP; P03299; IPOV.
 DR MEROPS; C03.001; -;
 DR MEROPS; C03.020; -;
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR009003; Cys_ser_trypsin.
 DR InterPro; IPR000199; Pept_3C_picorn.
 DR InterPro; IPR000081; Peptidase_C3.
 DR InterPro; IPR003138; Pico_P2A.
 DR InterPro; IPR002527; Pico_P2B.
 DR InterPro; IPR001676; Rhv.
 DR InterPro; IPR000605; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00548; Cys-protease-3C; 1.
 DR Pfam; PF02226; Pico_P2A; 1.
 DR Pfam; PF00947; Pico_P2A; 1.
 DR Pfam; PF01552; Pico_P2B; 1.
 DR Pfam; PF00073; Rhv; 3.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PF00918; CALICIVIRUSNS.
 DR ProDom; PD001125; Cys_protease_3C; 1.
 DR ProDom; PD001306; Pico_P2A; 1.
 DR ProDom; PD001274; Pico_P2B; 1.
 DR SMART; SM00382; AAA; 1.
 DR PolyProtein; Coat protein; Core protein; Transferase;
 DR RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
 DR Lipoprotein.
 FT CHAIN 2 69 COAT PROTEIN VP4.
 FT CHAIN 70 340 COAT PROTEIN VP2.
 FT CHAIN 341 578 COAT PROTEIN VP3.
 FT CHAIN 579 879 COAT PROTEIN VP1.
 FT CHAIN 880 1028 PROTEASE 2A.
 FT CHAIN 1029 1125 CORE PROTEIN 2B.
 FT CHAIN 1126 1454 CORE PROTEIN 2C.
 FT CHAIN 1455 1541 CORE PROTEIN 3A.
 FT CHAIN 1542 1563 GENOME-LINKED PROTEIN VP6.
 FT CHAIN 1564 1746 PICORNAIN 3C.
 FT CHAIN 1747 2207 RNA-DIRECTED RNA POLYMERASE 3D.
 FT LIPID 2 2 N-myristoyl glycine (in host).
 FT ACT_SITE 1710 1710 PROTEASE (POTENTIAL).
 FT ACT_SITE 1724 1724 PROTEASE (POTENTIAL).
 FT SEQUENCE 2207 AA; 245829 MW; 2B1E2070B7D44F99 CRC64;

Query Match
 Best Local Similarity 20.1%; Pred. No. 0.023; Indels 112; Gaps 20;
 Matches 99; Conservative 71; Mismatches 210;

QY 22 VPEVNASDPLAMDVPAGSGSTAVATAGQVNDIPDIPINNFFVQAPQGEFTISPN-----NTP 76
 DB 365 IPEFDVTPPI-----DIPGEVRNMELEADITMPLNLTNRKNTMDMYRVELNDAHSDTP 421
 QY 77 GGVLFDLSGLPHNPLHL-----LSQMYNGVGVNVRIMLAGNAFTAGKIIVSCIPPG 131
 DB 422 ---ILCLSLSPASDPLRAHTMLGEILNYYTHWAGSLKFTFLFCGSMWATGKLLVSYAPPG 478
 QY 132 FGSNLTIAQATLPHFVIAVDRTLDPIEVLPEEDVRNVLFNNDNRQQT-----MR 181
 DB 479 ABAPK-SRKEAMLGTHVIMDGLQSSCTMVVWNTYRQTINDSFTEGGVSMFYQTR 537
 QY 182 LVCMLYTLRTGGTGDSEFVAGRVNMTCSDFNLF-----LVPPT 223

Db 538 VVVPISPTPRKMD-----ILGFVSAC--NDFSVRLRLRDTTHISQEAAMPQGLDLIEGV 587
 QY 224 VEKTRPFTPLNPLSLNSRAPLPISGMGI SPDNVQSVQFQNGRCTLGRLVGTTPVS 283
 Db 588 VEGVTRNALTPLTANNLPDTQS-----SGPAHSKETPALTAVETG---ATNPLVPSDITVQ 640
 QY 284 LSHVAKIRGTNGTV-----INLTDLGTTPPHFEGPAPIGFPDILGGCDWHIN--- 331
 Db 641 TRHVIQKRTRSESTVESFFARGACVAIIEDV-----NDAPTRKASLFSVWKITYKD 692
 QY 332 ----MTQFGHSSQTYDVTTPDTFVPHLGSQAANGISGNYIGVLSWTSPPSHPSGSQV 387
 Db 693 TVQLRRKLEFFTYSRFDMEF---TFVVTISNYTDANNHALNQVYQIMYI-PPGAP----- 743
 QY 388 DLWKIPNYGSSITEATHLAPSVYPGPGEVLPFEMSKIIPGAGYSLPCLLPQEIYISHL-- 445
 Db 744 ----IPGKWDNYTQWTSNPS-----VFYTYGAP-PARISVYVYGIANAYSHFYD 788
 QY 446 ASEQAPTVEAAA 457
 Db 789 GFAKVPLAGQAS 800

RESULT 12

POLG_HE701 STANDARD; PRT; 2194 AA.
 ID POLG_HE701
 AC P32537;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Coat protein VP4 (PIA); Coat protein VP2
 DE (PIB); Coat protein VP3 (PIC); Coat protein P2C; Core protein P2A; Genome-
 DE linked protein VP6 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)
 DE (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 OS Human enterovirus 70 (strain J670/71) (EV 70).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Enterovirus.
 OX NCBI_TaxID=31915;
 RL [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=91037960; PubMed=2172447;
 RA Ryan M.D., Jenkins O., Hughes P.J., Brown A., Knowles N.J., Booth D.,
 RA Minor P.D., Almond J.W.;
 RT "The complete nucleotide sequence of enterovirus type 70:
 RT relationships with other members of the picornaviridae.";
 RL J. Gen. Virol. 71:2291-2299(1990).
 CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
 CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
 CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the
 CC poliovirus polyprotein. In other picornavirus reactions Glu may be
 CC substituted for Gln, and Ser or Thr for Gly.
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA) (N).
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; D00820; BAA18891.1; -;
 DR PIR; A36253; GNNYET.
 DR HSP; P03300; IPOV.
 DR MEROPS; C03.00A; -;


```

854 995 CORE PROTEIN P2A.
FT CHAIN 1090 CORE PROTEIN P2B.
1091 1412 CORE PROTEIN P2C.
1413 1489 CORE PROTEIN P3A.
FT CHAIN 1510 GENOME-LINKED PROTEIN VPG.
1511 1693 PICORNAIN 3C.
FT CHAIN 1694 RNA-DIRECTED RNA POLYMERASE P3D.
2 2 N-mristoyl glycine (in host).
1657 1657 PROTEASE (POTENTIAL).
FT ACT SITE 1671 PROTEASE (POTENTIAL).
1671 1671 PROTEASE (POTENTIAL).
FT CONFLICT 547 KD -> NH (IN REF. 1).
27 30
FT STRAND 36
FT HELIX 38
FT TURN 80
FT TURN 81
FT STRAND 83
FT TURN 87
FT TURN 88
FT STRAND 89
FT STRAND 90
FT STRAND 94
FT STRAND 101
FT HELIX 102
FT TURN 103
FT STRAND 113
FT STRAND 115
FT STRAND 123
FT HELIX 126
FT TURN 129
FT STRAND 133
FT STRAND 138
FT TURN 142
FT STRAND 147
FT TURN 152
FT HELIX 153
FT TURN 155
FT HELIX 156
FT HELIX 157
FT STRAND 159
FT STRAND 168
FT TURN 185
FT STRAND 188
FT TURN 197
FT TURN 198
FT STRAND 199
FT TURN 204
FT TURN 207
FT HELIX 213
FT HELIX 219
FT STRAND 223
FT TURN 225
FT TURN 232
FT HELIX 239
FT TURN 242
FT STRAND 246
FT HELIX 248
FT STRAND 256
FT TURN 261
FT STRAND 266
FT STRAND 280
FT TURN 282
FT STRAND 285
FT STRAND 288
FT TURN 303
FT STRAND 307
FT TURN 338
FT TURN 345
FT STRAND 353
FT TURN 356
FT STRAND 369
FT STRAND 372
FT TURN 373
FT HELIX 374
FT TURN 378
FT STRAND 381
FT TURN 385
FT STRAND 388
FT TURN 389
FT HELIX 395
FT STRAND 399
FT TURN 408
FT STRAND 411
FT TURN 419
420

422 426
FT TURN 428
FT HELIX 433
FT TURN 434
FT STRAND 436
FT STRAND 440
FT STRAND 443
FT STRAND 449
FT TURN 453
FT STRAND 454
FT STRAND 456
FT STRAND 458
FT HELIX 474
FT TURN 479
FT STRAND 481
FT STRAND 486
FT STRAND 492
FT STRAND 497
FT TURN 506
FT TURN 513
FT STRAND 518
FT STRAND 523
FT STRAND 528
FT TURN 531
FT STRAND 532
FT STRAND 537
FT TURN 545
FT STRAND 548
FT STRAND 550
FT STRAND 554
FT STRAND 567
FT HELIX 570
FT TURN 579
FT STRAND 580
FT STRAND 584
FT STRAND 588
FT TURN 601
FT STRAND 602
FT STRAND 603
FT HELIX 605
FT HELIX 618
FT STRAND 624
FT STRAND 629
629

Query Match 4.9%; Score 138; DB 1; Length 2153;
Best Local Similarity 19.5%; Pred. No. 0.099;
Matches 111; Conservative 80; Mismatches 201; Indels 178; Gaps 25;

QY 17 GAGQLVPEVNASDPLAMPVAGSATAVATAGQVN-----PIDPMIINNFVQAPQGE--- 67
Db 339 GSGQFMTTDDMQSPCAL-PWYHPTKEIFIPGVEKNLIEMCQVDTLIPINSTQSNIGNVSM 397
QY 68 --FTISPNNTPGGVLPF---DLSLGPLNLPFLHLSQMYNGVGNVRVIMLAGNAFTAG 121
Db 398 YTVTLSPQTKLAEBIFAIKVDIASHPPLATTLIGEIASYFTHWTSLSRFSFMCGTANTTL 457
QY 122 KIIVSCIPPGFGSHNLTIQAATLFPPIHVIADVTLDPIEVPLEDRVNLVFNHNDNRNQTM 181
Db 458 KVLLEYTPGIGKPR-SRKEAMLGTHVVDVCL-----QSTVS 494
QY 182 LVCMLYTLRTGGGTGDSFVAGRVMTCPSPDNFLFLVPPT----- 223
Db 495 LVVPWISASQVRFTTPTDYSSAGYI-TC---WYQTNFVVPPTNTAEMLCFVSGCKDFC 550
QY 224 -----VEQKTRPFT-----LPNLPLS--SLSNSRAPLPISGMG--- 254
Db 551 LRMARDTDLHKQTGPITQNPVERYDEVINEVLVVPENINQSHPTTSNAAPVLDAAETGHT 610
QY 255 --ISPDNVQSVQFQNGRCCTLDGRLVGTTPVSLSHVAKIRGTSNGTVINLTDLGTPFHP 312
Db 611 NKIQEDTETRYVQSSQTL-----EMSVESFLGRSGCHESVLDIVD----- 654
QY 313 EGPAPIGPDPDLGGCDWHINMTQ-----FGHSSQTQVDVDTPTTFVPHLGSIOANGIG 365
Db 655 -----NYNDQSFTKWINLQEMAIQIRKRFEMFTVARPDSI---TWVB---SVAAKOGH 702
QY 366 SGNYIGVLVSWSPSPHSPSGSQVD--LWKIPYNGSSITEATHLAPSVYPGPGFVLPFMS 423
Db 703 IGHIVQMYVYV--PPGAPITPTRDDYAWQ-----SGTNAS-----VFMOH 740
QY 424 KIPGPGAYSLPCLLPQEVISHLASQAPTVGAALLHY--VDPDTGRTLGEBKAYPDGFL 481
Db 741 GQFPF-RFSLPFL-----SIASAYMYFDGYDGDGT-----YKSRYG 775
QY 482 TCVPNGASSGQQLPFGVVFVSWVSRYF 511
```

DR 776 TWTNDMCTLCRSIVTSQHLKVKVWTRI 805

RESULT 15

POLG SVDVU

ID POLG SVDVU STANDARD; PRT; 2185 AA.

AC P1390; Q84794; Q84795; Q84796; Q84797; Q84798; Q84799; Q84800;

AC Q84801; Q84802; Q84803; Q84804;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins

DE P2A TO P2C; P2A; Genome-linked protein VPG; Picornain 3C

DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D

DE (EC 2.7.7.48)]

OS Swine vesicular disease virus (strain UKG/27/72).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Enterovirus.

OX NCBI_TaxID=12077;

[1]

SEQUENCE FROM N.A.

RX MEDLINE=90364770; PubMed=2168111;

RA Seachurn P., Knowles N.J., McCauley J.W.;

RT "The complete nucleotide sequence of a pathogenic swine vesicular

RL disease virus.";

RL Virus Res. 16:255-274 (1990).

CC -1- FUNCTION: It is thought that the P2C protein attaches to vesicular

CC membranes and is associated with viral RNA synthesis.

CC -1- FUNCTION: P2C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN

CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.

CC -1- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the

CC poliovirus polypeptide. In other picornavirus reactions Glu may be

CC substituted for Gln, and Ser or Thr for Gly.

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +

CC [RNA](N).

CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,

CC each of which is composed of one copy each of proteins VP1, VP2,

CC VP3, and VP4.

CC -1- PMW: Specific enzymatic cleavages in vivo yield mature proteins.

CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

EMBL; X54521; CAA38377.1; ..

DR MEROPS; C03.011; ..

DR MEROPS; C03.022; ..

DR InterPro; IPR004004; Calici_pol_hel.

DR InterPro; IPR009003; Cys Ser trypsin.

DR InterPro; IPR000199; pep3_c_picorn.

DR InterPro; IPR000081; Peptidase_C3.

DR InterPro; IPR003138; Pico_P1A.

DR InterPro; IPR002527; Pico_P2B.

DR InterPro; IPR001676; Rhv.

DR InterPro; IPR000605; RNA helicase.

DR InterPro; IPR007095; RNA_pol_DS_PS.

DR InterPro; IPR001205; RNA_pol_P3D.

DR InterPro; IPR007094; RNA_pol_PSVir.

DR InterPro; IPR008975; Viral_cap_coat.

DR Pfam; PF00548; Cys-protease-3C; 1.

DR Pfam; PF02226; Pico_P1A; 1.

DR Pfam; PF00947; Pico_P2A; 1.

DR Pfam; PF01552; Pico_P2B; 1.

DR Pfam; PF00073; Rhv; 3.

DR Pfam; PF00680; RNA_dep_RNA_pol; 1.

DR Pfam; PF00910; RNA helicase; 1.

DR PRINTS; PR00918; CALICVIRUSNS.

ProDom; PD001125; Cys_protease_3C; 1.

ProDom; PD001306; Pico_P2A; 1.

ProDom; PD001274; Pico_P2B; 1.

KW Polyprotein; Coat protein; Core protein; Transferase;

KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;

KW Lipoprotein.

FT CHAIN 2 69 COAT PROTEIN VP4 (P1A).

FT CHAIN 70 330 COAT PROTEIN VP2 (P1B).

FT CHAIN 331 568 COAT PROTEIN VP3 (P1C).

FT CHAIN 569 851 COAT PROTEIN VP1 (P1D).

FT CHAIN 852 1001 CORE PROTEIN P2A (P2-3B).

FT CHAIN 1002 1100 CORE PROTEIN P2B (P2-5B).

FT CHAIN 1101 1429 CORE PROTEIN P2C (P2-X).

FT CHAIN 1430 1518 CORE PROTEIN P3A (P3-1B).

FT CHAIN 1519 1540 GENOME-LINKED PROTEIN VPG (P3B).

FT CHAIN 1541 1723 PICORNAIN 3C.

FT CHAIN 1724 2185 RNA-DEPENDENT RNA POLYMERASE P3D.

FT CHAIN 2 2 N-myristoyl glycine (in host) (By

FT LIPID 2 similarity).

FT ACT_SITE 1687 PROTEASE (POTENTIAL).

FT ACT_SITE 1701 PROTEASE (POTENTIAL).

FT ACT_SITE 1701 PROTEASE (POTENTIAL).

SQ SEQUENCE 2185 AA; 243363 MW; C9B103052934E1B8 CRC64;

Query Match 4.8%; Score 136.5; DB 1; Length 2185;

Best Local Similarity 20.3%; Pred. No. 0.13;

Matches 118; Conservative 65; Mismatches 214; Indels 185; Gaps 28;

QY 1 MMASKDQATSSVPGAGAGQVPEVNASDPLAMPVAGSSTAVATAGQVNPIDPWLNINP 60

DB 324 LRLAGKQGLTTL-STPGSNQFLTSDDFQSPAM-POFDVTPEMDIFQV-----NNL 373

QY 61 VQAPQGFSTSPNTPGVL-----FDLSLGP--LNPFL--HL 96

DB 374 MEIAEVDVVPVNNTEGKVNISBAYQIPVQSNFTNSQVFGFLTPGANSVLNRTLLGEI 433

QY 97 SQMVGWGVNVRVIMLAGNAFTAGKIIIVSCIPPGFSHNLTIAQTALPPHVADVRLD 156

DB 434 LNYAHGSGSIKLTFFECGSMATKFLLAYSPGAGAPT-TRKAMLGTHVIWDV--- 488

QY 157 PIEVPLEDRNVLPHNNDRNQOTMRLVCMLYTLPLRTGGTGDSPV-----ACRVMT-- 209

DB 489 -----GLQSSCVLCIPWIS--QTHRYVVMDEYTAGYITCWY 524

QY 210 -----PSPDFNLFVPTVEQKTRPFTLPNPLSLSN--SRAPLPISG----- 252

DB 525 QTNIVVPADAQSDCKILCFVSACNDFSV--MLKDTFFIKQDNFFQGPGEVMGRAIARV 582

QY 253 ---MGISPDNVQSV-----QFQNGRCTLDGRVLGTPVSLSHVAKIRGTSNGTVINLT-- 302

DB 583 ADTIGSGPVNGESIPALTAETGHTS---QVPSDTMTQTRHVKNYHSRSSTVENFLCRS 639

QY 303 -----ELDGTFFHPFEGPAPIGFPDLGGCDWHNNMTQFGH-----SSQTYD 344

DB 640 ACVFYTTYKNHSDGDNF-----AYWINTROVAQLRRKLEMTFYARFD 683

QY 345 VDTTPDTTFVPHLGSIQANGISGNYIGVLS---WVSPSPHSGSQVDLWKIPNTGSSITE 401

DB 684 LEE---TFV---ITSTQEQTVPVGQDAPVLTHQIMVVPFGP-----VPTKVNYSW 729

QY 402 ATHLAPSVY-----PP-----QFGEVLVFF---MSKIPGPGAYSPLCLLPQEIYSHL 445

DB 730 QTSNDSVFEWTEGSAAPRMSIPFIFIGNAYSMTFYDGMARFKQGTGYGISTL----- 780

QY 446 ASEQAFTVGEAALLHVVDPTGTILGEFKAY---PDGFLTCVP 485

DB 781 -----NNMGTLVNRHVNDGPGPIVSTVRIYFKPKHVKTWVP 817

Search completed: June 1, 2004, 13:47:10

Job time : 7.40268 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 46.9047 Seconds
(without alignments)
3289.030 Million cell updates/sec

Title: US-09-926-799-3

Perfect score: 2895

Sequence: 1 MMASKAPQSGADGASGAGQ.....QLKPVGTASTARSRLGVRRRI 546

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Jan04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2895	100.0	546	4	AAB49702 Small rou
2	2231.5	77.1	544	4	AAB49703 Small rou
3	2034.5	70.3	530	2	AAR50972 Norwalk v
4	2034.5	70.3	530	7	Adc72176 Norwalk v
5	2029.5	70.1	530	4	AAB49701 Small rou
6	1993.5	68.9	530	2	AAR57091 Small rou
7	1977.5	68.3	545	4	AAB49700 Small rou
8	1175	40.6	535	4	AAB49707 Small rou
9	1169.5	40.4	540	4	AAB49706 Small rou
10	1155.5	39.9	542	4	AAB49708 Small rou
11	1124.5	38.8	548	4	AAB49705 Small rou
12	1115.5	38.5	548	5	AAU91272 Norwalk v
13	1105	38.2	539	4	AAB49704 Small rou
14	1090.5	37.7	550	4	AAB49709 Small rou
15	967	33.4	541	4	AAB49710 Small rou
16	321	11.1	579	2	AAW08143 RHDV caps
17	290	10.0	669	4	AAB67461 Amino aci
18	272	9.4	668	4	AAB67462 Amino aci
19	268.5	9.3	671	4	AAU50107 Feline ca
20	268	9.3	547	4	AAU50108 Feline ca
21	267	9.2	623	4	AAB47044 Feline Ca
22	267	9.2	623	4	AAB47043 Feline Ca
23	267	9.2	668	2	AAR10686 Feline ca
24	267	9.2	668	4	AAE04304 Feline ca
25	256	8.8	622	4	AAB47045 Feline Ca

26	144	5.0	2206	2	AAR22210	Aar22210 True type
27	134	4.6	40	5	AAU91273	AAU91273 Norwalk v
28	132.5	4.6	1194	6	ABU18239	ABU18239 Protein e
29	125.5	4.3	2209	1	AAP20037	Aap20037 Sequence
30	123	4.2	2735	6	ABR41356	ABr41356 Human DIT
31	123	4.2	2971	7	ADC35084	Adc35084 Human bre
32	123	4.2	2972	4	AAB50363	AAB50363 Human SRC
33	123	4.2	3118	4	AAB50362	AAB50362 Human SRC
34	122.5	4.2	6310	6	ABU39869	ABu39869 Protein e
35	121	4.2	40	5	AAU91274	AAu91274 Norwalk v
36	120	4.1	3443	2	AAW84559	AAw84559 Polyproti
37	119	4.1	2164	1	AAP81045	Aap81045 Sequence
38	118	4.1	2971	3	ABAB41231	ABa41231 Human ORF
39	117	4.0	2164	1	AAP80131	Aap80131 Peptides
40	116.5	4.0	2150	1	AAP60045	AAP60045 Sequence
41	116.5	4.0	2150	1	AAP81048	AAP81048 Sequence
42	116.5	4.0	2150	2	AAU05127	AAr05127 Complete
43	116	4.0	2016	4	ABB63911	ABb63911 Drosophil
44	115.5	4.0	239	1	AAP81097	ABP81097 Sequence
45	115.5	4.0	1765	6	ABU20131	ABu20131 Protein e

ALIGNMENTS

RESULT 1

AAAB49702
ID AAB49702 standard; protein; 546 AA.

XX AAB49702;

AC AAB49702;

XX 04-APR-2001 (first entry)

XX Small round structured virus protein SEQ ID 3.

XX Small round structured virus; SRSV; food poisoning.

XX Small round structured virus.

XX WO200079280-A1.

XX 28-DEC-2000.

XX 22-JUN-2000; 2000WO-JP004095.

XX 22-JUN-1999; 99JP-00175928.

XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

XX (DENK-) DENKA SEIKEN KK.

XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

XX WPI; 2001-080848/09.

XX N-PSDB; AAF29143.

XX Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.

XX Claim 1; Page 45-47; 84pp; Japanese.

XX This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

XX Sequence 546 AA;

Query Match 100.0%; Score 2895; DB 4; Length 546;
Best Local Similarity 100.0%; Pred. No. 1.5e-257;

Matches 546; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMASKDAPQSDAGSAGAGQVPEVNTADPLPMEFVAGTTAVATAGQVNMIDPWVNNF 60
DB 1 MMASKDAPQSDAGSAGAGQVPEVNTADPLPMEFVAGTTAVATAGQVNMIDPWVNNF 60

QY 61 VQSPQGEFTISPNNTPGDIILFDLQGLPHNLPFLSHLSQMYNGVGNMVRILLAGNAFSA 120
DB 61 VQSPQGEFTISPNNTPGDIILFDLQGLPHNLPFLSHLSQMYNGVGNMVRILLAGNAFSA 120

QY 121 GKIIICVCPGPTSSSLTIAQATLFPFHVADVRLTEPIEMPLEDVNRVLYHTNDNQPTMR 180
DB 121 GKIIICVCPGPTSSSLTIAQATLFPFHVADVRLTEPIEMPLEDVNRVLYHTNDNQPTMR 180

QY 181 LVCMLYTLPLRTGGSGNSDSFVVGAVRLTAPSSDFSFLVLPPTIEOKTRAFVNPILQ 240
DB 181 LVCMLYTLPLRTGGSGNSDSFVVGAVRLTAPSSDFSFLVLPPTIEOKTRAFVNPILQ 240

QY 241 TLSNRPSPSLIQMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLFRVRKINQGAR 300
DB 241 TLSNRPSPSLIQMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLFRVRKINQGAR 300

QY 301 TLNLTEVDGKPFMAFDSAPVPGFDFGKCDWHMRIKSTPNTSSGDPMRSVSVQTNVQGF 360
DB 301 TLNLTEVDGKPFMAFDSAPVPGFDFGKCDWHMRIKSTPNTSSGDPMRSVSVQTNVQGF 360

QY 361 VPHLGSIQDFEVFNHPTGDIYGTIEWISQSTPPGTDINLWEIPDYGSSLSQAANLAPPV 420
DB 361 VPHLGSIQDFEVFNHPTGDIYGTIEWISQSTPPGTDINLWEIPDYGSSLSQAANLAPPV 420

QY 421 FPPGFGALVVFVSAPFPNNRSAPNDVPCLLPQEIYTHFVSEQATMGDAALLHYVDDP 480
DB 421 FPPGFGALVVFVSAPFPNNRSAPNDVPCLLPQEIYTHFVSEQATMGDAALLHYVDDP 480

QY 481 TNRNLGEFKLYPGGYLTCVPGNGVAGPQOLPLNGVFLFVSWSRFYOLKPVGTASTARSR 540
DB 481 TNRNLGEFKLYPGGYLTCVPGNGVAGPQOLPLNGVFLFVSWSRFYOLKPVGTASTARSR 540

QY 541 LGVRR 546
DB 541 LGVRR 546

RESULT 2
ID AAB49703
XX AAB49703 standard; protein; 544 AA.
AC AAB49703;
DT AAB49703;
XX 04-APR-2001 (first entry)
XX Small round structured virus protein SEQ ID 4.
XX Small round structured virus; SRSV; food poisoning.
XX Small round structured virus.
XX WO200079280-A1.
XX 28-DEC-2000.
XX 22-JUN-2000; 2000WO-JP004095.
XX 22-JUN-1999; 99JP-00175928.
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI; 2001-080848/09.
XX N-PSDB; AAF29144.
XX

PT Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX Claim 1; Page 47-49; 84pp; Japanese.
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks

Sequence 544 AA;
Query Match 77.1%; Score 2231.5; DB 4; Length 544;
Best Local Similarity 74.7%; Pred. No. 2.2e-196; Indels 1; Gaps 1;
Matches 407; Conservative 59; Mismatches 78;

QY 1 MMASKDAPQSDAGSAGAGQVPEVNTADPLPMEFVAGTTAVATAGQVNMIDPWVNNF 60
DB 1 MMASKDAPQSDAGSAGAGQVPEVNTADPLPMEFVAGTTAVATAGQVNMIDPWVNNF 60

QY 61 VQSPQGEFTISPNNTPGDIILFDLQGLPHNLPFLSHLSQMYNGVGNMVRILLAGNAFSA 120
DB 61 VQSPQGEFTISPNNTPGDIILFDLQGLPHNLPFLSHLSQMYNGVGNMVRILLAGNAFSA 120

QY 121 GKIIICVCPGPTSSSLTIAQATLFPFHVADVRLTEPIEMPLEDVNRVLYHTNDNQPTMR 180
DB 121 GKIIICVCPGPTSSSLTIAQATLFPFHVADVRLTEPIEMPLEDVNRVLYHTNDNQPTMR 180

QY 181 LVCMLYTLPLRTGGSGNSDSFVVGAVRLTAPSSDFSFLVLPPTIEOKTRAFVNPILQ 240
DB 181 LVCMLYTLPLRTGGSGNSDSFVVGAVRLTAPSSDFSFLVLPPTIEOKTRAFVNPILQ 240

QY 241 TLSNRPSPSLIQMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLFRVRKINQGAR 300
DB 241 TLSNRPSPSLIQMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLFRVRKINQGAR 300

QY 301 TLNLTEVDGKPFMAFDSAPVPGFDFGKCDWHMRIKSTPNTSSGDPMRSVSVQTNVQGF 360
DB 301 TLNLTEVDGKPFMAFDSAPVPGFDFGKCDWHMRIKSTPNTSSGDPMRSVSVQTNVQGF 360


QY 361 VPHLGSIQDFEVFNHPTGDIYGTIEWISQSTPPGTDINLWEIPDYGSSLSQAANLAPPV 420
DB 361 VPHLGSIQDFEVFNHPTGDIYGTIEWISQSTPPGTDINLWEIPDYGSSLSQAANLAPPV 420

QY 421 FPPGFGALVVFVSAPFPNNRSAPNDVPCLLPQEIYTHFVSEQATMGDAALLHYVDDP 480
DB 421 FPPGFGALVVFVSAPFPNNRSAPNDVPCLLPQEIYTHFVSEQATMGDAALLHYVDDP 480

QY 481 TNRNLGEFKLYPGGYLTCVPGNGVAGPQOLPLNGVFLFVSWSRFYOLKPVGTASTARSR 540
DB 481 TNRNLGEFKLYPGGYLTCVPGNGVAGPQOLPLNGVFLFVSWSRFYOLKPVGTASTARSR 540

QY 541 LGVRR 545
DB 541 LGVRR 544

RESULT 3
AAR50972
ID AAR50972 standard; protein; 530 AA.
XX AAR50972;
XX AAR50972;
XX 16-OCT-2003 (revised)
XX 25-MAR-2003 (revised)
XX 05-OCT-1994 (first entry)
XX Norwalk virus strain 8P1a protein (encoded by ORF2).
XX

KW Norwalk virus; pathogen; acute gastroenteritis; food poisoning;
KW seafood contamination; diagnostic assay; calicivirus; small round virus.
OS Norwalk virus; (strain 8FIIa).
XX WO9405700-A2.
PN 17-MAR-1994.
PD 07-SEP-1993; 93WO-US008447.
XX 07-SEP-1992; 92US-00941365.
XX (BAYU) BAYLOR COLLEGE MEDICINE.
PA  Graham DY;
PI WPI: 1994-101125/12.
XX N-PSDB; AAQ56826.
DR DNA from Norwalk and related viruses - used for preparing prods. for use
PT in diagnostic assays, detection and vaccines for Norwalk and related
PT viruses.
XX Claim 14; Page 68-70; 156pp; English.
XX The Norwalk virus was isolated from stool samples from adult volunteers
CC infected with safety tested Norwalk virus strain 8FIIa. The coding
CC sequence is useful for the design of probes for use in diagnostic assays
CC for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct
CC PN field.) (Updated on 16-OCT-2003 to standardise OS field)
XX Sequence 530 AA;

Query Match 70.3%; Score 2034.5; DB 2; Length 530;

Best Local Similarity 69.4%; Pred. No. 3.1e-178; Mismatches 81; Indels 17; Gaps 6;

Matches 379; Conservative 69; Mismatches 81; Indels 17; Gaps 6;

1 MMASKADPQSDAGSAGQGVPEVNTADPLPMEPVAGTTAVATAGVNMIDPWINNF 60

1 MMASKDATSSVDGASGAGQGVPEVNASDPLAMDVPAGSTAVATAGVNPIDPWINNF 60

61 VOSPOGETTISPNNTPGDILFDLQGLPHNPLPFLSHLSQMYNGVGNMVRILLAGNAPSA 120

61 VQAPQGETTISPNNTPGDVLFDLSGLPHNPLPFLSHLSQMYNGVGNMVRIMLAGNAFTA 120

121 GKIIIVCCVPPGFTSSSLTIAQATLPHVIAVTRTLEPIEMPLEDVNLYHTND-NQPTM 179

121 GKIIIVSCIPPGGSHNLTAQATLPHVIAVTRTLEPIEMPLEDVNLYHTND-NQPTM 180

180 RLVCMLYTPLRGGSGNSDSFVAGRVLTAPSDPSFLVPPPTIEOKTRAFVTPNPL 239

181 RLVCMLYTPLRGGTG--DSFVAGRVMTCPSPDFNPLVPPVTEQKTRFTLPNPL 238

240 QTLNSRFPPLIQMILSPDASQVVOFQNGRCLIDGQLLGTTPATSGOLFVRGKINQGA 299

239 SSLNSRAPLPISSMIGSPDNVQSVQFQNGRCLIDGQLLGTTPATSGOLFVRGKINQGA 296

300 RLNLNTEVDGKPFMAFDSPAPVGFDFGKCDWMHRIKTPNTSGDPMRSVSUOTNVQ 359

297 TVINLTLDGTTFHFFEPGAPVGFDFGKCDWMHRIKTPNTSGDPMRSVSUOTNVQ 351

360 FVPHLGSIQFDFVFNHPTGDIYGTIEWISQPTPEGTIDINLWEIPDYSSISOANLAPP 419

352 FVPHLGSIQANGI---GSGNYGVLSWISPPSHPSGSGVDLWKIIPNYGSSITEATHLAPS 408

420 VPPPGFGEALVYFSAFPGPNRNSAPNDVPCLLPOEYITHFVSEQATMGDAALHYDDP 479

409 VYPPGFGVLFVFMKMPGP---GAYNLPCLLPQEYISHLASEQAPTVEAALLHYDDP 464

480 DTRNLGEBFKLYPGGLYTCVFNQVAGAGQOOLPLNGVFLFVSWSRFYQLKPVGTASTARS 539

465 DTGRNLGEBFKAYPDGFLTCVFNQVAGAGSGQQLPINGVFNVSWSRFYQLKPVGTASSARG 524

QY 540 RLQVRR 545
DB 525 RLGLRR 530

RESULT 4

ADC72176

ID ADC72176 standard; protein; 530 AA.

XX ADC72176;

XX ADC72176;

DT 18-DEC-2003 (first entry)

XX Norwalk virus protein 2 amino acid sequence.

DE immune response; non-Norwalk virus agent; immunogen; Norwalk virus;
KW viral pathogen; acute gastroenteritis; virucidal; antiviral vaccine.
XX Norwalk virus.

OS US6572862-B1.

XX 03-JUN-2003.

XX 07-JUN-1995; 95US-00486049.

XX 08-NOV-1989; 89US-00433492.

PR 27-APR-1990; 90US-00515993.

PR 27-AUG-1990; 90US-00573509.

PR 06-MAY-1991; 91US-00696454.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX  Graham DY;

XX WPI: 2003-776005/73.

DR N-PSDB; ADC72174.

Inducing an immune response against non-Norwalk virus agents, comprises administering an immunogen recombinantly expressed from a cDNA from Norwalk virus.
Example 4; SEQ ID NO 3; 45pp; English.
This invention relates to a novel method of inducing an immune response in an individual against Norwalk virus and non-Norwalk virus agents, by orally or parenterally administering an immunogen recombinantly expressed or synthesised from a cDNA of Norwalk virus given in the specification.
Norwalk virus is one of the most important viral pathogens, causing acute gastroenteritis. The invention may be used for the development of compounds with virucidal activity or an antiviral vaccine. The present sequence is the amino acid sequence of a protein encoded by the Norwalk virus genome of the invention.

Query Match 70.3%; Score 2034.5; DB 7; Length 530;

Best Local Similarity 69.4%; Pred. No. 3.1e-178; Mismatches 81; Indels 17; Gaps 6;

Matches 379; Conservative 69; Mismatches 81; Indels 17; Gaps 6;

1 MMASKADPQSDAGSAGQGVPEVNTADPLPMEPVAGTTAVATAGVNMIDPWINNF 60

1 MMASKDATSSVDGASGAGQGVPEVNASDPLAMDVPAGSTAVATAGVNPIDPWINNF 60

61 VOSPOGETTISPNNTPGDILFDLQGLPHNPLPFLSHLSQMYNGVGNMVRILLAGNAPSA 120

61 VQAPQGETTISPNNTPGDVLFDLSGLPHNPLPFLSHLSQMYNGVGNMVRIMLAGNAFTA 120

121 GKIIIVCCVPPGFTSSSLTIAQATLPHVIAVTRTLEPIEMPLEDVNLYHTND-NQPTM 179

121 GKIIIVSCIPPGGSHNLTAQATLPHVIAVTRTLEPIEMPLEDVNLYHTND-NQPTM 180

180 RLVCMLYTPRLTGGSGNSDSFVWAGRVLTAPSSDFSLFVLPPTIEQKTRAFVPMNPL 239
181 RLVCMLYTPRLTGGSGTG--DSFVWAGRVMTCPSPDFSLFVLPPTVEQKTRPFTLNLPL 238
240 QTLNSRFPSTLQGMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLFVRGKINQGA 299
239 SSLNSRAPLPISSMIGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGTSN--G 296
300 RTLNLTVDGKPFMAFDSPAPVFPDFGKCDWMHMRISKTPNNTSSGDPMRSVSVQTNVQ 359
297 TVINLTDLGTTPFHPPFEGPAPIGFDLGGCDWHINMTQFGHSSQT-----QYDVTTPDT 351
360 FVPHLGSIQDEVFENHPTDYIGTIEMISQPTTPGTDINLWIEIPDYGSSLSQAANLAPP 419
352 FVPHLGSIQANGI---GSGNYVGLVSWSPSPHSGSQVDLWKIPNYGSSITEATHLAPS 408
420 VPPGFGELVYFVSAPFPNRRSAPNDVPCLLPQEIYTHFVSEQAPTMGDAALLHYVDP 479
409 VPPGFGELVYFVMSKMPG---GAYNLPCLLPQEIYTHFVSEQAPTMGDAALLHYVDP 464
480 DTNRNLGEFKLYPGGYLTCVPGNGVAGPQPLNGVFLFVSVWSRFPYQLKPVGTASTARS 539
465 DTGRNLGEFKAYPDGFLTCVPGNGASSGPQQLPINGVVFVSVWSRFPYQLKPVGTASSARG 524
540 RLGVRR 545
525 RLGLRR 530

RESULT 5
AAB49701
ID AAB49701 standard; protein; 530 AA.
XX
AC AAB49701;
DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 2.
XX
KW Small round structured virus; SRSV; food poisoning.
OS Small round structured virus.
XX
FN WO200079280-A1.
XX
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000WO-JP004095.
XX
PR 22-JUN-1999; 99JP-00175928.
XX
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI; 2001-080848/09.
XX N-PSDB; AAF29142.
XX
PT Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
PS Claim 1; Page 42-45; 84pp; Japanese.
XX
CC This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAB20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX

SQ Sequence 530 AA;
Query Match 70.1%; Score 2029.5; DB 4; Length 530;
Best Local Similarity 69.4%; Pred. No. 8.9e-178;
Matches 379; Conservative 68; Mismatches 82; Indels 17; Gaps 6;
QY 1 MMASKADAPQSDAGSAGAGQLVPEYNTADPLPMEVAGPTTAVATAGQVNMIDPMVNNF 60
Db 1 MMASKADATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSTAVATAGQVNPIDPMIINN 60
QY 61 VQSPQGEFTISPNTPGDLFDLQGLPHNLPHSLSQMYNGWGNMVRVILLAGNAFSA 120
Db 61 VQAPQGEFTISPNTPGDLFDLQGLPHNLPHSLSQMYNGWGNMVRVILLAGNAFSA 120
QY 121 GKIIVCCVPPGFTSSSLTIAQAATLFPHVIAVDTLEPTLEMPLEDRVNLVYHTND--NQPTM 179
Db 121 GKIIVCSIPPGFSGHNLTTAQAATLFPHVIAVDTLEPTLEMPLEDRVNLVYHTND--NQPTM 180
QY 180 RLVCMLYTPRLTGGSGNSDSFVWAGRVLTAPSSDFSLFVLPPTIEQKTRAFVPMNPL 239
Db 181 RLVCMLYTPRLTGGSGTG--DSFVWAGRVMTCPSPDFSLFVLPPTVEQKTRPFTLNLPL 238
QY 240 QTLNSRFPSTLQGMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLFVRGKINQGA 299
Db 239 SSLNSRAPLPISSMIGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGTSN--G 296
QY 300 RTLNLTVDGKPFMAFDSPAPVFPDFGKCDWMHMRISKTPNNTSSGDPMRSVSVQTNVQ 359
Db 297 TVINLTDLGTTPFHPPFEGPAPIGFDLGGCDWHINMTQFGHSSQT-----QYDVTTPDT 351
QY 360 FVPHLGSIQDEVFENHPTDYIGTIEMISQPTTPGTDINLWIEIPDYGSSLSQAANLAPP 419
Db 352 FVPHLGSIQANGI---GSGNYVGLVSWSPSPHSGSQVDLWKIPNYGSSITEATHLAPS 408
QY 420 VPPGFGELVYFVSAPFPNRRSAPNDVPCLLPQEIYTHFVSEQAPTMGDAALLHYVDP 479
Db 409 VPPGFGELVYFVMSKMPG---LPCLLPQEIYTHFVSEQAPTMGDAALLHYVDP 464
QY 480 DTNRNLGEFKLYPGGYLTCVPGNGVAGPQPLNGVFLFVSVWSRFPYQLKPVGTASTARS 539
Db 465 DTGRNLGEFKAYPDGFLTCVPGNGASSGPQQLPINGVVFVSVWSRFPYQLKPVGTASSARG 524
QY 540 RLGVRR 545
Db 525 RLGLRR 530

RESULT 6
AAR57091
ID AAR57091 standard; protein; 530 AA.
XX
AC AAR57091;
XX
DT 27-AUG-2003 (revised)
XX 25-MAR-2003 (revised)
DT 05-OCT-1994 (first entry)
XX
DE Small round virus SRSV/KY/89 capsid protein.
XX pathogen; acute gastroenteritis; food poisoning; seafood contamination;
XX diagnostic assay; human calcivirus; small round virus; SRSV; KY89;
XX Norwalk virus; capsid protein.
XX
OS Small round structured virus.
XX
FN WO9405700-A2.
XX
PD 17-MAR-1994.
XX
PF 07-SEP-1993; 93WO-US008447.
XX
PR 07-SEP-1992; 92US-00941365.
XX

PA (BAYU) BAYLOR COLLEGE MEDICINE.
 XX Matson DO, Estes MK, Jiang X, Graham DY;
 XX
 XX WPI; 1994-101125/12.
 DR N-PSDB; AAQ56832.
 XX
 PT DNA from Norwalk and related viruses - used for preparing prods. for use
 PT in diagnostic assays, detection and vaccines for Norwalk and related
 PT viruses.
 XX
 XX Example 7; Fig 13a; 156pp; English.
 XX
 CC The known sequence for Norwalk virus was used to obtain the sequence of
 CC other Norwalk-related viruses such as SRSV/KY/89, an agent from a stool
 CC from an outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide
 CC cDNA sequence includes part of the polymerase region and the capsid
 CC region of the genome; the deduced amino acid sequences are AAR57092 and
 CC AAR57091, respectively. Expression of fragments and derivs. of Norwalk-
 CC related viruses permits development of diagnostic assays to detect
 CC antibodies, antigens, viral genetic material or antivirals. (Updated on
 CC 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS
 CC field.)
 XX
 XX Sequence 530 AA;
 SQ

Query Match 68.9%; Score 1993.5; DB 2; Length 530;
 Best Local Similarity 68.3%; Pred. No. 1.8e-174;
 Matches 373; Conservative 68; Mismatches 88; Indels 17; Gaps 6;

QY 1 MMASKADAPQADGASGAGQLVPEVNTADPLPMEPVAGPTTAVATAGQVNNIDPWVNNF 60
 DB 1 MMASKDATSSVDGASQVLVPEVNASDPLADPVDGASSTAVATAGQVNNIDPWVNNF 60
 QY 61 VOSPOGETTISPNNTPGDILFDLQGLPHLNPFLSHLSOMYNGWGNMVRILLAGNAFSA 120
 DB 61 VOAPQGETTISPNNTPGDILFDLQGLPHLNPFLSHLSOMYNGWGNMVRILLAGNAFTA 120
 QY 121 GKIIIVCCVPPGFTSSSLTIAQATLPPHVIADVRLTLEPIEMPLEDVRNLYHTND-NQPTM 179
 DB 121 GKIIIVCIPPGFGSQQLTIAQATLPPHVIADVRLTLEPIEMPLEDVRNLYHTND-NQPTM 180
 QY 180 RLVCMLYTLRTGGSGNSDSFVAVAGRVLTAPSSDFSLFLVPPPIEOKTRAFTVNNPL 239
 DB 181 RLVCMLYTLRTGGSGTG--DSFVAVAGRVLTAPSSDFSLFLVPPPIEOKTRAFTVNNPL 238
 QY 240 QTLNSRPFSLIQGMILSPDASOVVQFONGRCLIDQLGTTTPATSGQLFRVRGKINQGA 299
 DB 239 SLSNSRAPLPISGMIGISPDNVSQVQFONGRCLIDQLGTTTPATSGQLFRVRGKINQGA 296
 QY 300 RTLANLTVGDKPFMAFDPAPVGFDPFGKCDWHMIRISKTPTNNTSSGDPMSRSVSQTNVQ 359
 DB 297 TVINLTLDGTPPFHFEFGAPIGFDLGGCDWHMIRISKTPTNNTSSGDPMSRSVSQTNVQ 351
 QY 360 FVPHLGTSQFDEVFNHPTGDIYGTIEWTSQSTPTPGTDINWEIPDYGSSLSQAANLAPP 419
 DB 352 SVPHLGTSQANGI--GSGNYIGVLSVSPSPHPSGQVDLWKINYGSSITEATHLAPS 408
 QY 420 VFPFGGALYFVSAFPGPNRSAPNDVPCLLPOEYITHFVSEQAPTMGDAALHYDDP 479
 DB 409 VYSPFGGVLVFFMSKIFPGGDS----LPCLLPQGYISHLASEQAPTVGEGPLHYDDP 464
 QY 480 DTRNLGFKLYPGGYLTCVNGVAGPQQLPLNGVFLFVSVRSFYQLKPVGTASTARS 539
 DB 465 DTRNLGFKKAYPDGFLTCVNGAGSSGQQLPFGVFFVSVRSFYQLKPVGTASTARG 524
 QY 540 RLGVRR 545
 DB 525 RLGLRR 530

RESULT 7
 AAB49700

ID AAB49700 standard; protein; 545 AA.
 XX
 AC AAB49700;
 XX
 DT 04-APR-2001 (first entry)
 XX
 DE Small round structured virus protein SEQ ID 1.
 XX
 KW Small round structured virus; SRSV; food poisoning.
 XX
 OS Small round structured virus.
 XX
 PN WO200079280-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 22-JUN-2000; 2000WO-JP004095.
 XX
 PR 22-JUN-1999; 99JP-00175928.
 XX
 PA (NTNA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.
 XX
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 XX
 DR WPI; 2001-080848/09.
 DR N-PSDB; AAF29141.
 XX
 PT Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX
 PS Claim 1; Page 40-42; 84pp; Japanese.
 XX
 CC This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX
 SQ Sequence 545 AA;

Query Match 68.3%; Score 1977.5; DB 4; Length 545;
 Best Local Similarity 66.6%; Pred. No. 5.8e-173;
 Matches 367; Conservative 76; Mismatches 95; Indels 13; Gaps 7;

QY 1 MMASKADAPQADGASGAGQLVPEVNTADPLPMEPVAGPTTAVATAGQVNNIDPWVNNF 60
 DB 1 MMASKDAPTNDGTSGAGQLVPEANTAEPTSMPEVAGATAAATAGQVNNIDPWIMNY 60
 QY 61 VOSPOGETTISPNNTPGDILFDLQGLPHLNPFLSHLSOMYNGWGNMVRILLAGNAFSA 120
 DB 61 VOAPQGETTISPNNTPGDILFDLQGLPHLNPFLSHLSOMYNGWGNMVRILLAGNAFTA 120
 QY 121 GKIIIVCCVPPGFTSSSLTIAQATLPPHVIADVRLTLEPIEMPLEDVRNLYHTNDNQPTMR 180
 DB 121 GKIIIVCIPPGFAQNTSIAQATLPPHVIADVRLTLEPIEMPLEDVRNLYHTNDNQPTMR 180
 QY 181 LVCMLYTLRTGGSGNSDSFVAVAGRVLTAPSSDFSLFLVPPPIEOKTRAFTVNNPLQ 240
 DB 181 LVCMLYTLRASGSSGTDPPVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSVNNPLN 240
 QY 241 TLSNRPFSLIQGMILSPDASOVVQFONGRCLIDQLGTTTPATSGQLFRVRGKI--NOG 298
 DB 241 TLSNRVPSLISKMWVRSDHGMQVQFNGRVTLTDLQGLGTTTPATSGQLFRVRGKI--NOG 298
 QY 299 ARTLNTEVDGKPFMAFDPAPVGFDPFGKCDWHMIRISKTPTNNTSSGDPMSRSVSQTNVQ 358
 DB 301 GNGYNLTGDSYPHAFESPAPIGFPDLGECDDHWEASPT-TQFNTGDIYKQINVKQES-S 358
 QY 359 GFVPHLGSIQFDEVFNHPTG-DYIGTIEWISQSTPTPGTDINLWEIPDYGSSLSQAANLA 417

359 AFAPHLGTIQADGLSDVSNTNMIAKLGWSPVSDGHRGDVDPWVPIRYGSTLLEAAQLA 418
 418 PPVFPFGGEALVYFVSAPP---GNNSAPNDVPCLLPOEYITHFYSEQAPTMGDAALL 474
 419 PPIYPGGEATVFFMSDFPIAHGTNGLS---VPTTIPQBFVTHFVNEQAPTRGEAALL 474
 475 HVVDPTNRNLGEEFKLYPGGYLTCVPGVAGPQQLPLNGVFLFVSWSRFYQLKPVGTA 534
 475 HYLDDPTNRNLGEEFKLYPGGYLTCVPGVAGPQQLPLNGVFLFVSWSRFYQLKPVGTA 534
 535 STARSRLGVRR 545
 535 GPA-CRLGIRR 544

RESULT 8
 AAB49706
 ID AAB49706 standard; protein; 535 AA.

XX
 AC AAB49706;
 DT 04-APR-2001 (first entry)
 DE Small round structured virus protein SEQ ID 8.
 DE Small round structured virus; SRSV; food poisoning.
 KW Small round structured virus.
 OS Small round structured virus.
 PN WO200079280-A1.
 PD 28-DEC-2000.
 PF 22-JUN-2000; 2000WO-JP004095.
 PR 22-JUN-1999; 99JP-00175928.
 PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 XX WPI; 2001-080848/09.
 DR N-PSDB; AAF29148.
 XX Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX Claim 1; Page 57-59; 84pp; Japanese.
 PS This invention relates to a kit for the detection and typing of small
 XX round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX
 XX Sequence 535 AA;
 SQ Query Match 40.6%; Score 1175; DB 4; Length 535;
 Best Local Similarity 46.2%; Pred. No. 5.2e-99;
 Matches 261; Conservative 82; Mismatches 172; Indels 50; Gaps 16;
 1 MMASNDAPQSDAGSAGQOLVPEVNTADPLPMEPVAGETTAATAGQVNMIDPWINVF 60
 1 MKVASNDAPSDNDGAG---LVPEANN-ETMALEPVAGASIAAPLTGQNNIIDPWIRLNF 56
 61 VQSPQGEFTISNPTGDIILFDLQGLHNPFLSHLSOMYNGVGNMVRILLAGNATSA 120
 57 VQAPNGEFTVSPRNSPGEVLNLELGLPELNPYLAHLNRYINGAGGVEVQVLLAGNATFA 116

121 GKIIVCCVPPGTSSTSLTIAQATLPHPHVIADVRLTEPIEMPLEDVRNLYHTN-DNQPTM 179
 117 GKLVAFAVPPHPHLEINSPGQITMPEPHVIDVRLTEPLVLLPDPVRNNFFHYNQNEPRM 176
 180 RLVMCLYTLRTTGGSGNSDSFVWAGRVLTAPSSDFSFLVPPPTIEQKTRAFVNIPL 239
 177 RLVMCLYTLPLRS-NGSGD-DVFTVSCRVLTRSPDFDNYLVPPTLESKTKFTLPIITI 234
 240 QTLNSRFRPSLIQCMTLSPASQVQFONGRCLIDGQLLGTTPATSGQLFRVRGKINQGA 299
 235 GELTNSRFPVPIDELYTSPNESLVQFQNGRCALDGELOQTTLQLLPTAICSFGRINQKV 294
 300 R-----TLNLTBVDGKPF-MAFDSAPVGFDP-CKCDMHRIRISKTPNNTSSGDDPMRSV 351
 295 SGENHWNMQVTNINGTPDPDTPGVPAPLGTDFSGKLFGLVLSQRDHDNACRSHDAV--- 351
 352 SVQTNVQGFVPHLGSIOFDEVNHPITGDIYGTIE-----WISOPS---TPPG---TDINLW 401
 352 -IATNSAKFTPLGAIQ-----IGTWEEDDVHINQPTKFTPVGLFENEGFNQW 398
 402 EIPDYSSLSQAANLAPPVPPGFGALVYFVSAPFPNNRSAPNDVPCLLPOEYITHFV 461
 399 TLPNYSGALTNNMGLAPPVAPTPFGQILFFRSHIPLKGGVADP-VIDCLLPQEWIQLY 457
 462 SEQAPTMGDAALLHYVDPTNRNLGEPKLYPGGYLTCVPGVAGPQQLPLNGVFLFVSW 521
 458 QESAPSQSDVALIRFTNPTDGRVLFEAKLRSGYITVA--NTGSRPIVVPANGFFRFDTW 515
 522 VSRFYQLKPVGTASTARSRLGVRR 546
 516 VQFYSLAPMGTGN-----GRRRV 534

RESULT 9
 AAB49706
 ID AAB49706 standard; protein; 540 AA.

XX
 AC AAB49706;
 DT 04-APR-2001 (first entry)
 DE Small round structured virus protein SEQ ID 7.
 DE Small round structured virus; SRSV; food poisoning.
 KW Small round structured virus.
 OS Small round structured virus.
 PN WO200079280-A1.
 PD 28-DEC-2000.
 PF 22-JUN-2000; 2000WO-JP004095.
 PR 22-JUN-1999; 99JP-00175928.
 XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 XX (DENK-) DENKA SEIKEN KK.
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 XX WPI; 2001-080848/09.
 DR N-PSDB; AAF29147.
 XX Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX Claim 1; Page 54-57; 84pp; Japanese.
 PS This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC

CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX
 SQ Sequence 540 AA;

Query Match 40.4%; Score 1169.5; DB 4; Length 540;
 Best Local Similarity 44.9%; Pred. No. 1.7e-98;
 Matches 254; Conservative 87; Mismatches 178; Indels 47; Gaps 13;
 QY 1 MMASKDAPQSDAGSAGQOLVPEVTADPLPMEPVAGPTTAVATAGVNMIDPMVNF 60
 DB 1 MMASNDATPSNDGAG--LVPESSN-EAMALEPVVGASLAAPVTGQNIIDPMIRNF 56
 QY 61 VQSPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLSQMYNGWGNMVRILLAGNAFSA 120
 DB 57 VOAPNGEFTVSPRNSPGEILNLELGPENLPYLAHARMYNGYAGMEVQVNLGNAFTA 116
 QY 121 GKIIIVCCVPPGPTSSSLTIAQATLPHVIAVRLTLEPIEMPLEDVRNVLHTND-NQPTM 179
 DB 117 GKIIIFAAVPPYPVENLSPSQITMPEPHVIDVRLTLEPVLPMVDVRLTLEHFNQKDEPKM 176
 QY 180 RLVCMLYPLRTGGSGNSDSFVAGRVLTAPSSDFSLFLVPTIEQKTRAFVTPNIP 239
 DB 177 RLVALMYLTPLRN-NGSGD-DVFTVSCRILTRESPEFDLYLPPTVESKTKPFTLPVLT 234
 QY 240 QTLNSRPPSLTQGMILSPDASOVVQFONGRCLIDGQLLGTTPATSGOLFRVRGKI---- 295
 DB 235 GELNSRPPSLTQGMILSPDASOVVQFONGRCLIDGQLLGTTPATSGOLFRVRGKI---- 295
 QY 296 --NQARTLNLTVEVGKPFMAFDS-PAPVGFPPDFGKCDWHMRISKTPNNTSSGDPMSVS 352
 DB 295 PSEOHMNLITNLNGTQDPTDDVPALGVDFAGEVGLVSQRNGESNPANRAHDAV 354
 QY 353 VQTNVQGVPHLGSQF-----DEVNHTDGYIGTIEWISOPSTPGTD-----INL 400
 DB 355 VATVSDKTPKLGVLQIGTWNVDVENQFT-----KFTPIGLNEVANGHRFEQ 402
 QY 401 WEIPDYGSSLSQAAMLAPPVFPFGGEALVYFVSAPFGPNRNSAPNDVPCLLPQBYITHF 460
 DB 403 WTLPRYSALTLNMLNLAFAVAPLTFGERLLLFRRSVYPLKGGFNP-AIDCSVPQEWVHF 461
 QY 461 VSEQAPTMGDAALHYVDPTNRLNGLBFKLYPGGYLTCPVNGVAGPQOLPLNGVLFVS 520
 DB 462 VQESAPSLGDVALRVYNPDGTGRVLFKALHKGGLTV--SSTSTGTPVVPVANGYFKFDS 519
 QY 521 WVSRYQLKPVGTASTARSRLGVRI 546
 DB 520 WVNQFYSLAPMGTCN-----GRRRV 539

RESULT 10

AAAB49708
 ID AAAB49708 standard; protein; 542 AA.

XX AAAB49708;

AC 04-APR-2001 (first entry)

DT Small round structured virus protein SEQ ID 9.

DE Small round structured virus; SRSV; food poisoning.

XX Small round structured virus.

OS WO200079280-A1.

XX 28-DEC-2000.

XX 22-JUN-2000; 2000WO-JP004095.

XX 22-JUN-1999; 99JP-00175928.

XX
 PA
 XX

(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 (DENK-) DENKA SEIKEN KK.

Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

WPI: 2001-080848/09.
 N-PSDB; AAF29149.

Kit for the detection and typing of small round-structured virus (SRSV)
 strains for investigation of food poisoning outbreaks, contains
 antibodies.

Claim 1; Page 59-61; 84pp; Japanese.

This invention relates to a kit for the detection and typing of small
 round structured virus (SRSV) strains. The kit contains antibodies
 directed against peptides represented in sequences AAB49700 - AAB49710,
 which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 used for detecting and typing strains of SRSV in order to prevent the
 spread of infection and to examine the epidemiology of outbreaks

XX Sequence 542 AA;

Query Match 39.9%; Score 1155.5; DB 4; Length 542;
 Best Local Similarity 45.3%; Pred. No. 3.4e-97;
 Matches 258; Conservative 88; Mismatches 171; Indels 53; Gaps 18;

QY 1 MMASKDAPQSDAGSAGQOLVPE-VNTADPLPMEPVAGPTTAVATAGVNMIDPMVNF 59
 DB 1 MMASNDATPSNDGAG--SLVPEGIN--ETMPLEPVAGASIAAPVAGQNTIIDPMIRNF 55
 QY 60 FVQSPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLSQMYNGWGNMVRILLAGNAFSA 119
 DB 56 FVQAPNGEFTVSPRNSPGEILNLELGPENLPYLAHARMYNGYAGMEVQVNLGNAFTA 115
 QY 120 AKGIIVCCVPPGPTSSSLTIAQATLPHVIAVRLTLEPIEMPLEDVRNVLHTNDN-QPT 178
 DB 116 AGKILFAAIPNPNFLVDMISPAQITMLPHLIVDVRLTLEPIEMPLEDVRNVLHTNDN-QPT 175
 QY 179 MRLVCMYLTPLRTGGSGNSDSFVAGRVLTAPSSDFSLFLVPTIEQKTRAFVTPNIP 238
 DB 176 MRLVAMLYTPLRS-NGSGD-DVFTVSCRILTRESPEFDLYLPPTVESKTKPFTLPILT 233
 QY 239 LQTLNSRPPSLTQGMILSPDASOVVQFONGRCLIDGQLLGTTPATSGOLPRVRGKI--- 295
 DB 234 ISELNRSRPPPIEQLYTAPNETNVVQCNGRCLIDGQLLGTTPATSGOLPRVRGKI--- 293
 QY 296 --NQARTLNLTVEVGKPFMAFDS-PAPVGFPPDFGKCDWHMRISKTPNNTSSGDP--MR 349
 DB 294 NGDNWDQNLQLQTYPNNGASYDPTDEVPAPLGTQDFSGMLYGV-LTQDNVNVSTGEAKNAK 352
 QY 350 SVSVQTNVQGVPHLGSQFDEVFN--HPTGDYIGTIEWISQPS--TPPG-----TDIN 399
 DB 353 GIYISTTSKTKPKIGSLGHSITEHVHP-----NQSRFTPVGVAVENTPFG 401
 QY 400 LWEIPDYGSSLSQAAMLAPPVFPFGGEALVYFVSAPP--GPNRNSAPNDVPCLLPQBY 456
 DB 402 QWVLPFYAGSLALNTNLAPAVAPTTPGQQLFFRRSVPCVQLOQSDA--FIDCLLPQEW 459
 QY 457 ITHFVSEQAPTMGDAALHYVDPTNRLNGLBFKLYPGGYLTCPVNGVAGPQOLPLNGV 516
 DB 460 VNHFYQEAAPQADVALRVYNPDGTGRVLFKALHRSQFIIV--SHTGAYPLVVPVNGHF 517
 QY 517 LFVSNVSRFYOLKPVGTASTARSRLGVRI 546
 DB 518 RFDVWNVQFYSLAPMGTCN-----GRRRI 541

RESULT 11

AAAB49705
 ID AAAB49705 standard; protein; 548 AA.

XX AAB49705;
AC 04-APR-2001 (first entry)
DT Small round structured virus protein SEQ ID 6.
DE Small round structured virus; SRSV; food poisoning.
KW Small round structured virus.
OS Small round structured virus.
XX WO200079280-A1.
XX 28-DEC-2000.
XX 22-JUN-2000; 2000WO-JP004095.
XX 22-JUN-1999; 99JP-00175928.
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI; 2001-080848/09.
DR N-PSDB; AAF29146.
XX Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX Claim 1; Page 52-54; 84pp; Japanese.
XX This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX Sequence 548 AA;
SQ
Query Match 38.8%; Score 1124.5; DB 4; Length 548;
Best Local Similarity 43.0%; Pred. No. 2.5e-94; Indels 69; Gaps 17;
Matches 250; Conservative 80; Mismatches 182;
QY 1 MMASKDAPQADGAGAGQGVVNTADPLPMEPVAGPTTAVATAGQVNNIDPWVNNF 60
DB 1 MMASNDAPNDGAG--LVPEINN-EAMALDPVAGAAATAAALPTGQNIIDPWVNNF 56
QY 61 VQSGQGEFTISPNPTGDIPLDQLGPHLNPFLSHLSQMYNGWVGNVRVILLAGNAFSA 120
DB 57 VQAPGGEFTVSPRNSPGEVLLNLELGPENINPYLAHARMYNGYAGGFEVQVVLGNAFTA 116
QY 121 GKIIIVCCVPPGFTSSSLITTAQATLFPVHIVADVRLTEPLEMPLDVRNVLYHTNDNQPT-M 179
DB 117 GKIIIFAAIPPNPFIIDNLGAAQITMCPHVIVDVRQLPEVNPVLPMPDVRNFFHYNQSURL 176
QY 180 RLVCMLYPLRTGGSGNSDSFVAVAGVLTAPSSDFSLFVPPTEIOKTRAFVPIPL 239
DB 177 RLIIAMLYPLR--ANNSGDVFTVSCRVLTRPSDFSNFLVPPVTPVESKTKPFTLPILT 234
QY 240 QTLNSRPFSLIQGMILSPASQVQFQNGRCLIDGQLLGTTPATSGQLFRVGRKINGOA 299
DB 235 SEMSNRRFPVPIESLRTENIVVQCGRVTLDGELMTGTLQLLPSQICAFRGVLTST 294
QY 300 -----RTLNL-----LTEVDGKPF-MAFDSAPVGPDPF-GKCDWHMRISKTP 339
DB 295 SRASDAQDTATPLFNYHVVQLDNLNGTFYDPAEDIPGLTGPDPFRGK---VFGVASQ 350
QY 340 NNTSSGDMRVSQVNVQVFPVHIGSIQ-----FDEVFNHPTGDIYGTIEWISQPSST 392

351 RNLDSTTRAHEAKVDTTAGRTFKLGSLEISTSDDFDQ--NQPT-----KFT 396
393 PFG-----TDINLWEIPDYSSLSQAANLAPPVPPGFGALVYFVSAPGPNRSAPN 446
397 PVGIGVDNEABFQWLSLFDYSGQFTNNMLAPAVAFNPFGBQLLFFRSQLPSSGGRS--N 454
447 DV-PCLLPQEIYTHFVSEQAPTMGDAALHLYVDDPTNRNLGFEKLYPGGYLTCTVNGVGA 505
455 GVLDCLVPQEWQHFQYESAPAQTOVALVRYVNDTGVLPKALHKLGFMTIANN--GD 512
506 GPOQLPLNGVFLFVSVWSRYQLKPVGTASTARSRLGVRI 546
513 SPITVPNGVFRFESWVNDFTLAPMTGN-----GRRRI 547

RESULT 12
AAU91272
ID AAU91272 standard; protein; 548 AA.
XX AAU91272;
AC AAU91272;
DT 18-JUN-2002 (first entry)
XX Norwalk virus associated polynucleotide #1.
DE Norwalk virus; monoclonal antibody; geno group I; geno group II;
KW immunological detection; food; viral infection.
XX Norwalk virus.
OS JP2002020399-A.
XX 23-JAN-2002.
XX 10-JUL-2000; 2000JP-00208151.
XX 10-JUL-2000; 2000JP-00208151.
XX (OSAP) OSAKA PREFECTURE.
PA (NISE-) NIPPON SEIBUTSU KAGAKU CENT KK.
PA (IATR) IATRON LAB INC.
XX WPI; 2002-287412/33.
XX A monoclonal antibody useful in the immunological detection and diagnosis
PT of Norwalk virus infection.
XX Disclosure; Page 12-13; 24pp; Japanese.
XX The invention describes a monoclonal antibody recognising Norwalk virus,
CC a capsid protein of Norwalk virus, or a common antigen epitope on the
CC capsid protein molecule of geno group I and geno group II. The antibody
CC is useful for immunological detection and quantitative analysis of
CC Norwalk virus in foods and the serum of infected patients. This sequence
CC represents a Norwalk virus associated protein described in the invention
XX Sequence 548 AA;
SQ
Query Match 38.5%; Score 1115.5; DB 5; Length 548;
Best Local Similarity 42.7%; Pred. No. 1.7e-93; Indels 61; Gaps 15;
Matches 243; Conservative 83; Mismatches 182;
QY 1 MMASKDAPQADGAGAGQGVVNTADPLPMEPVAGPTTAVATAGQVNNIDPWVNNF 60
DB 1 MMASNDAPNDGAG--LVPEINN-EAMALDPVAGAAATAAALPTGQNIIDPWVNNF 56
QY 61 VQSGQGEFTISPNPTGDIPLDQLGPHLNPFLSHLSQMYNGWVGNVRVILLAGNAFSA 120
DB 57 VQAPGGEFTVSPRNSPGEVLLNLELGPENINPYLAHARMYNGYAGGFEVQVVLGNAFTA 116
QY 121 GKIIIVCCVPPGFTSSSLITTAQATLFPVHIVADVRLTEPLEMPLDVRNVLYHTNDNQPT-M 179
DB 117 GKIIIFAAIPPNPFIIDNLGAAQITMCPHVIVDVRQLPEVNPVLPMPDVRNFFHYNQSURL 176

Qy	180	RLVCLYTLRLTGGSGNSDSFVAVAGRVLTAPSDSFSLFLVPPPTIEQKTRAFVTPNPL	239
Db	177	RLIAMLYTLR--ANNSGDDVFTVSCRVLTPSPDPFSFNFLVPPTVESKTKPFTLPILT	234
Qy	240	QTLNSRPFPSITQGMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQGA	299
Db	235	SEMNSRFPVPPIESLHTSPTENIVQCQNGRVTLDGELMGTTQLLPSRICAFRGVLTST	294
Qy	300	-----RTLN-----LTEVDGKPF-MAFDSDPAPGVGDFGKCDWMRISKTN	340
Db	295	SRASDAQDIATPLRFNYHYHVLQDLNLGTPYDAEDI PGPLGTDPFRGKVGVSQRNPD	354
Qy	341	NTSSGDPMRSVQTNVQGVFPHLGSIQ-----FDEVFNHTGDYIGTIEWISQSPSTP	393
Db	355	STTRA--HEAKVDTTAGRFPPKLSLEISTESSDFDQ--NQPT-----RFTP	397
Qy	394	PG-----TDINLWEIPDYGSSLSQAANLAPVPFPFGFGEALYFVSAPFGPNNSRAPND	447
Db	398	VGIGVDNEADPQWSLPDYSGOFTNNMLAPAVAPNFPGEQLLFFRSQLPSSGGRS--NG	455
Qy	448	V-PCLLPOEYITHFVSECAPTMGDAAALAHYVDPDTNRNLGEEKLYPGGVLTCTVPGNVGAG	506
Db	456	VLDCLVPOEWQHFQYESAPAQTVQALVRYVNPDTGRVLFEAKLHKLGEMTIAKN--GDS	513
Qy	507	PQQLPLNGVFLFVSVWSRYQLKPVGTAS	535
Db	514	PITVPPNGYFRFESWVNPFFYTLAPMGTN	542
RESULT	13		
ID	AAB49704		
XX	AAB49704 standard; protein; 539 AA.		
AC	XX		
AC	XX		
XX	AAB49704;		
DT	XX		
DT	04-APR-2001 (first entry)		
XX	XX		
DE	XX		
DE	Small round structured virus protein SEQ ID 5.		
XX	XX		
KW	XX		
KW	Small round structured virus; SRSV; food poisoning.		
XX	XX		
OS	XX		
OS	Small round structured virus.		
XX	XX		
PN	XX		
PN	WO200079280-A1.		
XX	XX		
PD	28-DEC-2000.		

22-JUN-2000; 2000WO-JP004095.
22-JUN-1999; 99JP-00175928.
(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
(DENK-) DENKA SEIKEN KK.
Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
WPI; 2001-080848/09.
N-FSDB; AAF29145.
Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.
Claim 1; Page 50-52; 84pp; Japanese.
This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

XX	Sequence 539 AA;
SQ	
	Query Match 38.2%; Score 1105; DB 4; Length 539; Best Local Similarity 44.4%; Pred. No. 1.5e-92; Matches 248; Conservative 80; Mismatches 194; Indels 36; Gaps 17;
Qy	1 MMASDKADPQASGAGSAGOLVPEVNTADPLPHEPVAGPTAVATACQVMIDWIYNNF 60 Db 1 MKMASNDANPS--DGST--ANKLVPEVNN-EYMALPEVVGAIAAPAVAGQQNVDPWIRNF 56 Qy 61 VQSPQGFTTSPNNTPGDILFDLQLGPHLNPFSLHSOMYNGVMGNRVILLIAGNAFSA 120 Db 57 VQAPGGFTVSPRNAGEILLWSAPLGPDLPYLISHLARMTNGVAGGEVQVILLAGNAFTA 116 Qy 121 GKIIIVCCPPGFTSSSIITIAQAATLFPHVIADVRLTEPIEMPLEDRNLVYHTN-DNQPTM 179 Db 117 GKIIIFAAVPNPFTTEGLSPSQVTMPHHIIVDVQRLEPVLPLDVRNNFYHNSNDSTI 176 Qy 180 RLVCMLYTPURTCGGSGNSDSFVAGRVLTAPSDFSFLFLVPTTBQKTRAFTVPNIP 239 Db 177 KLIALMYTPLR--ANNAGDDVFVTSRCVLRTRSPDFDI FLVPPPTVESRTKPTVILT 234 Qy 240 QTLNSRPFPSLIOGMILSPDASOVQFQGRCLIDGLGTTTATSGQLFRVRGKINQGA 299 Db 235 EEMSNGRFPIPLEKLYTGPSA FVQFQNGRCTTDGVLLGTTLQSAYNICTFERGDVTHIA 294 Qy 300 ----RTLNLTEVDGKPFMAPDS-PAPVGFPDF-GKCDWHMRIKTPNNTSSGDPMRSVS 353 Db 295 GSHDYTMNLASQWNSVDYTEEIPAPLCTFD FVKIQG--MLQTTFREDGS-TRAHKATV 351 Qy 354 QTNVQGFVPHLSIGFDEVFNH--PTGDY-----IGTIWISQPSTPPGTDINLWEIPDY 406 Db 352 STGSVHFPTKLGSVQVYTTDTNDFOTGNTKFTFGVGIQDGNHNQNEP----QQWLPLNY 407 Qy 407 GSLSQMANLAPVFPFGFEALVYFSAPFPNNRSAPN-DVPCLLPQBYITHFVSEQA 465 Db 408 SGRTHNVHLAPAVAFTFPGEQLLFPRSTRMPGCS--GYPMNLDCLLPQSWVQHFCQEA 465 Qy 466 PTMGDAALLHYVDPTNRNLGEFKLYPGGYLTCPVNGVGAGPQOL--PLNGVFLFVSWS 523 Db 466 PAQSDVALLRFVNPTDGRVLFECKLHKSQVTV----AHTGPHDLVIPPNGYFRFDSWN 521 Qy 524 RPYQLKFPVGTASTARSRL 541 Db 522 QFTVTLAPMGNGAGRRAAL 539
RESULT 14	AAB49709 standard; protein; 550 AA.
XX	AAB49709;
AC	AAB49709;
XX	
DT	04-APR-2001 (first entry)
XX	
DE	Small round structured virus protein SEQ ID 10.
XX	
KW	Small round structured virus; SRSV; food poisoning.
XX	
OS	Small round structured virus.
XX	
PN	WC200079280-A1.
PD	28-DEC-2000.
XX	
PF	22-JUN-2000; 2000WO-JP004095.
XX	
PR	22-JUN-1999; 99JP-00175928.
XX	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES. (DENK-) DENKA SEIKEN KK.
XX	

Wed Jun 2 09:13:27 2004

us-09-926-799-3.rag

```

PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI; 2001-080848/09.
XX DR N-PSDB; AAF29150.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX PT strains for investigation of food poisoning outbreaks, contains
XX PT antibodies.
XX
XX PS Claim 1; Page 62-64; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
XX CC round structured virus (SRSV) strains. The kit contains antibodies
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX CC used for detecting and typing strains of SRSV in order to prevent the
XX CC spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 550 AA;
XX
XX Query Match 37.7%; Score 1090.5; DB 4; Length 550;
XX Best Local Similarity 42.8%; Pred. No. 3.4e-91;
XX Matches 250; Conservative 75; Mismatches 173; Indels 89; Gaps 16;
XX
QY 1 MMASKDAPOSADGAGAGQVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPWINNF 60
DB 1 MKWASNDAPSDNGAA---NLVPEAND-EVMALEPVVGASIAAPVVGQNIIDPWINNF 56
QY 61 VQSPQGEFTTSPNNTPGDILOLQPHLNPFLSHLSQMYNGWGNMVRILLAGNAPSA 120
DB 57 VQAPQGEFTVSPRNSPGEMLNLELGPDLNPYLHSLRMNGVAGGQVQVVLGNAFTA 116
QY 121 GKIIIVCCVPPGFTSSSLTIAQATLPHVIAVDTLEPIEMPLEDVRNLYHTN-DNQPTM 179
DB 117 GKIIIFAAVPPHFPVENISAQITMCPHVIIVDVRQLEVPVLLPLDIRNPFHYNQENTPRM 176
QY 180 RLVCMLYTPRLTGGSGNSDSFVAVGRVLTAPSSDFSFLVPPTIEOKTRAFTVNPIL 239
DB 177 RLVCMLYTPLRANSQ---EDVFTVSCRVLTRPAPDEFTFLVPTVESKTKPFTLPILT 233
QY 240 QTLNSRFPSSLIQMILSPDASQVVFQNGRCILDGQLLGTTPATSGQLFRVRGK-INQG 298
DB 234 GELNSRFPAAIDMLYTDPNESIIVVQPNQGRCTLDGLTQGTQTLVPTQICAFRGTLSQT 293
QY 299 ARTLNLYE-----VDGKPFMAFDS-PAPVGFDPFGKCDWHMRISKTPNN 341
DB 294 ARAADSTDSQORARNHPLHVQVKNLDGTYDPTDDIPAVLGAIDF-----KG 340
QY 342 TSSGDPMRSVQTNVQ-----FVPHLSIQF-----DEVFNHP 376
DB 341 TVFG-----VASQRVQSQEQGHVYATRAHAHIDTDPKYPKLGTLILKSGDDFN-- 393
QY 377 TGDYIGTIEWISQ--STPPGTDIN---LWEIPDYGSSLSQAANLAPPVPPPGFEALVY 431
DB 394 -----THQPIRFTFVGNDNNWQWELPDYSGRLTLMNMLAPAVSPSPGERILF 443
QY 432 FVSAPFPGNRSAPNDVCLLPQEVYTHFVSEQAPTMDGDAALLHYVDPDTNRLGEFKLY 491
DB 444 FRSIVPSAGGYG--GYIDCLIPQEWVQHFYQEAAPSAVALRVYVNDPTGRNIFEAKHL 502
QY 492 PGGYLTCVPNGVAGPQQLPLNGVFLFVSWVSREYQKLPVGTASTAR 538
DB 503 REGFLTVA--NCGNNPIVPPNGYPRFEAWGNQFTLAPMSGQGR 547
XX
XX RESULT 15
XX ID AAB49710
XX AC AAB49710 standard; protein; 541 AA.
XX AC AAB49710;
XX
XX DT 04-APR-2001 (first entry)

```

```

XX Small round structured virus protein SEQ ID 11.
XX
XX Small round structured virus; SRSV; food poisoning.
XX
XX Small round structured virus.
XX
XX WO200079280-A1.
XX
XX 28-DEC-2000.
XX
XX 22-JUN-2000; 2000WO-JP004095.
XX
XX 22-JUN-1999; 99JP-00175928.
XX
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX PA (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI; 2001-080848/09.
XX DR N-PSDB; AAF29151.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX PT strains for investigation of food poisoning outbreaks, contains
XX PT antibodies.
XX
XX PS Claim 1; Page 64-66; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
XX CC round structured virus (SRSV) strains. The kit contains antibodies
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX CC used for detecting and typing strains of SRSV in order to prevent the
XX CC spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 541 AA;
XX
XX Query Match 33.4%; Score 967; DB 4; Length 541;
XX Best Local Similarity 38.5%; Pred. No. 7.9e-80;
XX Matches 225; Conservative 88; Mismatches 188; Indels 84; Gaps 17;
XX
QY 1 MMASKDAPOSADGAGAGQVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPWINNF 60
DB 1 MKWASNDAPSDNGAAG---LVPEINN-EVMPLEFVAGASLATPVVGQNIIDPWINNF 56
QY 61 VQSPQGEFTTSPNNTPGDILOLQPHLNPFLSHLSQMYNGWGNMVRILLAGNAPSA 120
DB 57 VQAPQGEFTVSPRNSPGELLDLELGPDLNPYLHSLRMNGVAGGMEVQIVLAGNAFTA 116
QY 121 GKIIIVCCVPPGFTSSSLTIAQATLPHVIAVDTLEPIEMPLEDVRNLYHTN-DNQPTM 179
DB 117 GKIIIFAAVPPGPPYENLSPSQITMCPHVIIDVRQLEPFLPMPDIIWNNFFHYNQNDPKL 176
QY 180 RLVCMLYTPRLTGGSGNSDSFVAVGRVLTAPSSDFSFLVPPTIEOKTRAFTVNPIL 239
DB 177 RLVCMLYTPLR---ANNSGDDVFTVSCRVLTKRSPDFEFTFLVPTVESKTKQFALPILKI 234
QY 240 QTLNSRFPSSLIQMILSPDASQVVFQNGRCILDGQLLGTTPATSGQLFRVRGKI---N 296
DB 235 SEMTNSRFPVVDVVMYTARNENQVVPQNGRVTLDGELLGTLTFLLAIVNICKFKEVIAKN 294
QY 297 QGART-----LNLTEVDGKPF-MAFDSAPVGPDPFGKCDWHMRISKTPNNTSSGDPMRSV 351
DB 295 GDVRSYRMDMEITNTDGTIDPTEDTDPGIPGSPDFQGLFGVASQRNKNKEQNPAATRAHEA 354
QY 352 SVCTNVQGFVPHLSIQFDEVF-----NHTGDIYIG 382
DB 355 IINTGGDLHCQISS---SEIYLTSPNLRCTNPPQLPQSGRLGTLILRSDNGHCHDMVG 411
QY 383 TIEWISQPTTPTGTDINLMEIPDYGSSLSQAANLAP-PVFPFGFEALVYFVSAPPGPN 441

```


Db 412 T-----SPTTP--TWQOWRCRSGSNCCSGHRYVPVMMNRVTWIVLSHKSGFSTSTR 464
QY 442 RSAPNDVPCLLPQEIYTHFVSEQAPTMGDAALLHYVDPDTNRNLGEFKLYPGGYLTCVEN 501
Db 465 K-----LQ-----LNLWP-----LIRFINPDTGRVLFPEARLHKQGITVA-- 501
QY 502 GVGAGPQQLPLNGVFLFVSWVSFFYQLKPVGTASTARSLGVRI 546
Db 502 HTGDNPIVMPPNGYRFRFEAWNQFYSLAPVGTGK-----GRRRV 540

Search completed: June 1, 2004, 13:45:48
Job time : 48.9047 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 13.0087 Seconds
(without alignments)
2166.837 Million cell updates/sec

Title: US-09-926-799-3
Perfect score: 2895
Sequence: 1 MMASKDAPQSDAGSAGAGQ.....QLKPVGTASTARSRLGVRR1 546

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS-COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2034.5	70.3	530	US-08-486-049-3	Sequence 3, Appli
2	290	10.0	669	US-09-617-594A-2	Sequence 2, Appli
3	272	9.4	668	US-09-617-594A-4	Sequence 4, Appli
4	268.5	9.3	626	US-09-590-020-7	Sequence 7, Appli
5	267	9.2	623	US-09-590-020-2	Sequence 2, Appli
6	267	9.2	623	US-09-590-020-4	Sequence 4, Appli
7	256	8.8	622	US-09-590-020-6	Sequence 6, Appli
8	140	4.8	2206	US-07-852-260-2	Sequence 2, Appli
9	140	4.8	2206	US-08-461-503-2	Sequence 2, Appli
10	140	4.8	2206	US-08-465-250-2	Sequence 2, Appli
11	123	4.2	2972	US-09-579-181-2	Sequence 2, Appli
12	123	4.2	3118	US-09-579-181-1	Sequence 1, Appli
13	120	4.1	3443	US-08-416-603-2	Sequence 2, Appli
14	109.5	3.8	2736	US-09-252-991A-30227	Sequence 30227, A
15	109	3.8	2322	US-09-976-594-15	Sequence 15, Appl
16	107	3.7	1048	US-09-171-699-10	Sequence 10, Appl
17	106	3.7	2318	US-09-091-219-24	Sequence 24, Appl
18	106	3.7	2318	US-09-660-541-24	Sequence 24, Appl
19	105.5	3.6	580	US-08-672-564-1	Sequence 1, Appli
20	102.5	3.5	1911	US-09-854-856-64	Sequence 64, Appl
21	102.5	3.5	1939	US-09-854-856-48	Sequence 48, Appl
22	102.5	3.5	1971	US-09-854-856-32	Sequence 32, Appl
23	102.5	3.5	1999	US-09-854-856-16	Sequence 16, Appl
24	102.5	3.5	2004	US-09-854-856-58	Sequence 58, Appl
25	102.5	3.5	2032	US-09-854-856-42	Sequence 42, Appl
26	102.5	3.5	2048	US-09-854-856-62	Sequence 62, Appl
27	102.5	3.5	2064	US-09-854-856-26	Sequence 26, Appl

28	102.5	3.5	2076	4	US-09-854-856-46	Sequence 46, Appl
29	102.5	3.5	2092	4	US-09-854-856-10	Sequence 10, Appl
30	102.5	3.5	2108	4	US-09-854-856-30	Sequence 30, Appl
31	102.5	3.5	2136	4	US-09-854-856-14	Sequence 14, Appl
32	102.5	3.5	2141	4	US-09-854-856-56	Sequence 56, Appl
33	102.5	3.5	2157	4	US-09-854-856-52	Sequence 52, Appl
34	102.5	3.5	2169	4	US-09-854-856-40	Sequence 40, Appl
35	102.5	3.5	2185	4	US-09-854-856-36	Sequence 36, Appl
36	102.5	3.5	2201	4	US-09-854-856-24	Sequence 20, Appl
37	102.5	3.5	2217	4	US-09-854-856-20	Sequence 20, Appl
38	102.5	3.5	2229	4	US-09-854-856-8	Sequence 8, Appl1
39	102.5	3.5	2245	4	US-09-854-856-4	Sequence 4, Appl1
40	102.5	3.5	2294	4	US-09-854-856-50	Sequence 50, Appl
41	102.5	3.5	2322	4	US-09-854-856-34	Sequence 34, Appl
42	102.5	3.5	2354	4	US-09-854-856-18	Sequence 18, Appl
43	102.5	3.5	2382	4	US-09-854-856-2	Sequence 2, Appl1
44	100.5	3.5	580	4	US-09-252-991A-23699	Sequence 23699, A
45	100.5	3.5	721	4	US-09-390-234-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-08-486-049-3
; Sequence 3, Application US/08486049
; Patent No. 6572862
; GENERAL INFORMATION:
; APPLICANT: BEIG, Mary K.
; APPLICANT: Guang, Xi
; APPLICANT: Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Ave., N.W.
; CITY: Washington, D.C.
; STATE:
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,049
; FILING DATE: June 7, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311.023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-486-049-3

Query Match 70.3%; Score 2034.5; DB 4; Length 530;
Best Local Similarity 69.4%; Pred. No. 1.9e-193;
Matches 379; Conservative 69; Mismatches 81; Indels 17; Gaps 6;

QY 1 MMASKDAPQSDAGSAGAGQLVPEVNTADPLPMEVAGPTTAVATAGQVNMIDPWVNF 60
DB 1 MMASKDATSSVDGASGAGQLVPEYNASDPLAMDPAVGSSTAVATAGQVNPIDPWVNF 60

QY 61 VSPQGEFTISNNTPGDIPLDQLGPHLNPFLSHLSOMYNGWGNMVRILLAGNAFSA 120
DB 61 VQAPQGEFTISNNTPGDIPLDQLGPHLNPFLSHLSOMYNGWGNMVRILLAGNAFTA 120
QY 121 GKIIIVCCVPPGPTSSSLTIAQATLPPHVIADVRLTEPIEMPLEDVRNVLVYHTND-NQPTM 179
DB 121 GKIIIVCICPPGSHNLITIAQATLPPHVIADVRLTEPIEMPLEDVRNVLVYHTND-NQPTM 180
QY 180 RLVCMLYTPRTGGSGNSDSFVAVAGRVLTAPSSDFSLFVPPPTIEOKTRAFVTPNIPL 239
DB 181 RLVCMLYTPRTGGSGT--DSFVAVAGRVMTCPSPDFNLFVPPPTIEOKTRAFVTPNIPL 238
QY 240 OTLSNRFSLIOGMLTSPDASOVQFQNGRCLIDGQLLGTTPATSGQLFRVRGKINOGA 299
DB 239 SSLSNRAPLPTSSMIGISPDNVQSVQFQNGRCLIDGRLVGTTPVSLSHVAKIRGTSN--G 296
QY 300 RTNLNTEVDGKPPMAFDSAPVGFDFGKCDWHMRISKTPNNTSSGDMRVSVSQTNVQG 359
DB 297 TWINLTDLGTPPHFEGPAPIGFDFLGGCDWHINNTQFHSQT-----QYDVTITPDT 351
QY 360 FVPHLSIQDFEVNHTPDYICTIEWISQSPSTPGTDINLWEIPDYGSSLSQAANLAPP 419
DB 352 FVPHLSIQANGI---GSGNYGVLSWISPPSHPSGSQVDLMKIPNYGSSITEATHLAPS 408
QY 420 VFPFGFGEALVYFVSAPFPPNRSAPNDVPCLLPQBYITHFVSEQAPTMGDAALLHYVDP 479
DB 409 VYPPGFGEVLVFFMSKMPG---GAYNLPCLLPQBYISHLASEQAPTMGEAALLHYVDP 464
QY 480 DTRNLGEFLKYPGCGYLTCPVNGVAGPQOLPLNGVFLFVSWVSRYQLKPVGTASTARS 539
DB 465 DTGRLGEFKAYPDGFLTCVPNGASSGQQLPVGPFVSVWSRYQLKPVGTASSARG 524
QY 540 RLGVRR 545
DB 525 RLGLRR 530

RESULT 2
US-09-617-594A-2
; Sequence 2, Application US/09617594A
; Patent No. 6541458
; GENERAL INFORMATION:
; APPLICANT: Audomnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.1
; CURRENT APPLICATION NUMBER: US/09/617,594A
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-617-594A-2

Query Match 10.0%; Score 290; DB 4; Length 669;
Best Local Similarity 25.5%; Pred. No. 9,8e-20;
Matches 153; Conservative 71; Mismatches 247; Indels 130; Gaps 29;

QY 12 ADGASGAGOLVPEVNT-ADPLPMEPVAGPTTAVATAGQVNMIDPVIWNVFVQSGEFTI 70
DB 125 ADGDSI--TTPEQGLVGVIAEAPSAQATAADATGKSDSEW--ESFF-SFTSVNW 179
QY 71 SPNNTGDIPLDQLGPHLNPFLSHLSOMYNGWGNMVRILLAGNAFSAKIIIVCCVPP 130
DB 180 STSETQKILFKQSLGPLNLPYLSKLYVAVSGSDVRFISGSGVFGKLAIVVPP 239

QY 131 GFTS--SSLTIAQATLPPHVIADVRLTEPIEMPLEDVRNVLVYHTNDNQFTMELVCMLYTPL 189
DB 240 GVDVFOVSTMLQ---YPHVLFDARQVEPVISIPDLRSTLYHLMSDQTTTSLIVMVYNDL 296
QY 190 -RTGGSGNSDSFVAVAGRVLTAPSSDFSLFVPPPTIEOKTRAFVTPNIPLQTLNS-- 245
DB 297 INPYANDSNSSGCIVT--VETKPGDPFKPHLLKPP-----GSMLTGHSIPSLIPKSSSL 349
QY 246 ----RPPSLIOGMLTSPDASOVVQFQNGRCLIDGQLLGTTPATSGQLFRVRGKINOGART 301
DB 350 WIGNRYMSDITDFVIRP-----FVFOANRHFDFNQ---ETAGWSTPRFR-----PIT 393
QY 302 LNLTEVDGKPP--MAFDSAPVGFDFGKCDWHMRISKTP-----NNTSSGDP 347
DB 394 ITLSNGSKLGTGVATDIYP-GIFD-GWPDITIGEELTPAGDYSITNGSGNDIATANA 451
QY 348 MRSVSQTNVQGVPHLGSIQDFEVNHTPDYI-CTIB--W-----ISOP 390
DB 452 YDSADVITNTNF-----RGMVICALQORAMGDKKISSTAITTAKEG 495
QY 391 ST-PPGTDINLWEIPD-----GSSLSQAANLAPPVFPFGFGEAL-----VYFVSAPP 437
DB 496 NTLKPSNTIDMTKIAVQDTHVGRDVTSDDTLAILGYTGIGEOAIGSNRDSVVRISMLP 555
QY 438 GPNNRSA-----PNDVPCLLPQBYITHFVSEQAPTMGDAALLHYVDP-----DT 481
DB 556 ETGARGGNHPIFYKNSIKLGVLSRSDVNSQILHTSRQLSLNHYLLPPDSFAVYRIIDS 615
QY 482 NRNLGEFLKYPGCGYLTCPVNGVAGPQ-QLPLNGVFLFVSWVSRYQLKPVGTASTARS 540
DB 616 NGSFDFVGDSDGSFV---GVSSIPKLEFPFLSASYMGI-----QLAKIRLASNIRST 665
QY 541 L 541
DB 666 M 666

RESULT 3
US-09-617-594A-4
; Sequence 4, Application US/09617594A
; Patent No. 6541458
; GENERAL INFORMATION:
; APPLICANT: Audomnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V
; FILE REFERENCE: 454313-3151.1
; CURRENT APPLICATION NUMBER: US/09/617,594A
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-617-594A-4

Query Match 9.4%; Score 272; DB 4; Length 668;
Best Local Similarity 28.7%; Pred. No. 6e-18;
Matches 99; Conservative 39; Mismatches 135; Indels 72; Gaps 15;

QY 21 LVEVNATA-DP-LPMSPVAGPTTAVATAGQVNMIDPVIWNVFVQSGEFTI----- 70
DB 106 LIGEVAKAMPNLPLFRLREADDGSVTTPEQGT-----LVGGVIAEAPNAQMSAVADVATCK 160
QY 71 -----SPNNTGDIPLDQLGPHLNPFLSHLSOMYNGWGNMVRILL 112
DB 161 SVDSWEAEPFSTSVNWTSETQKILFKQSLGPLNLPYLSKLYVAVSGSIEVRFS 220

QY 113 LAGNAPSAGKIIVCCVPPGFTS--SSLTIAQATLPHVIAVDTLPIEMPLEDVRNLVH 171
Db 221 ISGSGVFGKLAIAIVPPGIDPVQSTMLQ---YPHVLFDAQVPEVIFETIPDLRSLVH 277
QY 172 TNDNQPTMLVCMWLYTPL--RTGGSGNSDSFVVAGRVLTAPSDPSFLVPTIEQKTR 230
Db 278 LMSDITTSLSVIMYNDLINPYANDNSNGCIVT--VETKPGDFKFLHLKPP-----GS 330
QY 231 AFTVNPNIQTLNS-----RFPSLIQMILSPDASQVQFQNGRCLIDQGLLGTTPA 283
Db 331 MLTHGSIPLPKSSSLWGNRHWSDIIDFKIP-----FVFOANRHDFNQ---ETAG 382
QY 284 TSGQLFVRVGKINQAGARTLNLTVEGKPF---MAFDSAPVGPFD 325
Db 383 WSTPRFR-----PITTVSEKGGKLGIGVATDSIVP-GIPD 418

RESULT 4
US-09-590-020-7
; Sequence 7, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE REFERENCE: CALICIVIRUS AND VACCINES THEREOF
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 626
; TYPE: PRF
; ORGANISM: Feline calicivirus
US-09-590-020-7

Query Match 9.3%; Score 268.5; DB 4; Length 626;
Best Local Similarity 27.5%; Pred. No. 1.2e-17;
Matches 112; Conservative 45; Mismatches 153; Indels 97; Gaps 18;

QY 10 QSADGASGAGLVPEVNT--ADPLPMEPVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEF 68
Db 79 EADDGSGITA-----PEQGTWVGVIAPSPSQMSTADMATGKSVDSSEW-----EAPF 125
QY 69 TI-----SPNPTPGDILFDLQGLPHLPFLSHLSQMYNGVGNMRVIRILLAGNAPSAGK 122
Db 126 SPHTSVNWSSTETQKILFKQSLGPLNLPYLAKLYVWAGSSIEVRFSISGSGVFGK 185
QY 123 IIVCCVPPGFTS--SSLTIAQATLPHVIAVDTLPIEMPLEDVRNLVHTNDNQPTML 181
Db 186 LAAIVVPPGVDPVQSTMLQ---YPHVLFDAQVPEVIFCLPDLRSLVHLSMDITTS 242
QY 182 VCMWLYTPL--RTGGSGNSDSFVVAGRVLTAPSDPSFLVPTIEQKTRAEFTVNIPLQ 240
Db 243 VLMVYNDLINPYANDNSNGCIVT--VETKPGDFKFLHLKPP-----GSMITHGVSFSD 295
QY 241 TILNS-----RFPSLIQMILSPDASQVQFQNGRCLIDQGLLGTTPATSGQLFRVRG 293
Db 296 LIPKSSSLWGNRHWSDIIDFKIP-----FVFOANRHDFNQ---ETAGWSTPRFR--- 344
QY 294 KINQAGARTLNLTVEGKPF---MAFDSAPVGPFDGKCDWHRISKTPNTSSGDPKRS 350
Db 345 -----PISVTITEQNGAKLGIGVATDSIVP-GIPD-----GW-----PDTTIPGELI-- 385
QY 351 VSVQINVQGFVHLGSIQFDEVFNHTGDYIGTIEWISQSPSTPPGTD 397
Db 386 -----PAGDYAITNGTNDITATTGYD 407

RESULT 5
US-09-590-020-2
; Sequence 2, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE REFERENCE: CALICIVIRUS AND VACCINES THEREOF
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRF
; ORGANISM: Feline calicivirus
US-09-590-020-2

Query Match 9.2%; Score 267; DB 4; Length 623;
Best Local Similarity 31.3%; Pred. No. 1.7e-17;
Matches 73; Conservative 32; Mismatches 88; Indels 40; Gaps 8;

QY 21 LVPEVNTA-DP-LPMEPVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEFTI-----70
Db 61 LIARVAAEAWDDPLFLRLEGDDGSITTEQGTM-----VGVIAEPSSQMSAADMATOK 115
QY 71 -----SPNPTPGDILFDLQGLPHLPFLSHLSQMYNGVGNMRVIRILL 112
Db 116 SVDSSEWAEAFSPHTSVNWSSTETQKILFKQSLGPLNLPYLAKLYVWAGSSIEVRFS 175
QY 113 LAGNAPSAGKIIVCCVPPGFTS--SSLTIAQATLPHVIAVDTLPIEMPLEDVRNLVH 171
Db 176 ISGSGVFGKLAIAIVPPGIEPVQSTMLQ---YPHVLFDAQVPEVIFAIPLDRSLVH 232
QY 172 TNDNQPTMLVCMWLYTPL--RTGGSGNSDSFVVAGRVLTAPSDPSFLVPTIE 223
Db 233 LMSDITTSLSVIMYNDLINPYANDNSNGCIVT--VETKPGDFKFLHLKPP 283

RESULT 6
US-09-590-020-4
; Sequence 4, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE REFERENCE: CALICIVIRUS AND VACCINES THEREOF
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 623
; TYPE: PRF
; ORGANISM: Feline calicivirus
US-09-590-020-4

Query Match 9.2%; Score 267; DB 4; Length 623;
Best Local Similarity 31.3%; Pred. No. 1.7e-17;
Matches 73; Conservative 32; Mismatches 88; Indels 40; Gaps 8;

QY 21 LVPEVNTA-DP-LPMEPVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEFTI-----70

Db 61 LIAKVAEAWDPLPLFRLEGDDGSIITPEQGTM-----VGGVIAEPPSAQMSAAADMATGK 115
Qy 71 -----SPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWVGNMVRIL 112
Db 116 SVDSWEAFPSFHTSVNWSSTQKILFKQSLGPLNPNYLSHLAKLYVANSQVEVRF 175
Qy 113 LAGNAFSAKIIIVCCVPPGFTS--SSLTIAQATLFPFHVIAVRTLEPIEMPLEDVNRVLYH 171
Db 176 ISGSGVFGKLAIAIVVPPGIEPVQSTMLQ---YPHVLFDAQVPEVIFAIPDLRSNLYH 232
Qy 172 TNDNQPTMRLVCMLYTPL-RTGGSGNSDSFVAVAGVLTAPSSDSSELSFLVPP 223
Db 233 LMSDTDTTSLVIMVYNDLINPYANDTNSGCIVT--VETKPGDPKFKHLKPP 283

RESULT 7
US-09-590-020-6
; Sequence 6, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vainis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 622
; TYPE: PRN
; ORGANISM: Feline calicivirus
; US-09-590-020-6

Query Match 8.8%; Score 256; DB 4; Length 622;
Best Local Similarity 30.9%; Pred. No. 2.1e-16;
Matches 72; Conservative 33; Mismatches 89; Indels 40; Gaps 8;
Qy 21 LVEVNTA-DP-LPMEVAGPTTAVATAGQNMIDPWVNNFVQSQGEETI-----70
Db 61 LIAKVAEAWDPLPLFRLEGDDGSIITPEQGTM-----VGGVIAEPPSAQMSAAADMATGK 115
Qy 71 -----SPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWVGNMVRIL 112
Db 116 SVDSWEAFPSFHTSVNWSSTQKILFKQSLGPLNPNYLSHLAKLYVANSQVEVRF 175
Qy 113 LAGNAFSAKIIIVCCVPPGFTS--SSLTIAQATLFPFHVIAVRTLEPIEMPLEDVNRVLYH 171
Db 176 ISGSGVFGKLAIAIVVPPGIEPVQSTMLQ---YPHVLFDAQVPEVIFAIPDLRSNLYH 232
Qy 172 TNDNQPTMRLVCMLYTPL-RTGGSGNSDSFVAVAGVLTAPSSDSSELSFLVPP 223
Db 233 LMSDTDTTSLVIMVYNDLINPYANDTNSGCIVT--VETKPGDPKFKHLKPP 283

RESULT 8
US-07-852-260-2
; Sequence 2, Application US/07852260
; Patent No. 5525715
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham

STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/852,260
FILING DATE: 19920619
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 36607-B-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2206 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-852-260-2
Query Match 4.8%; Score 140; DB 1; Length 2206;
Best Local Similarity 21.2%; Pred. No. 0.0006;
Matches 122; Conservative 62; Mismatches 221; Indels 170; Gaps 27;
Qy 17 GAGQLVPEVNTADPLPMEPVAGPTTAVATAGV-NM-----IDPWVNNFVQSQG---66
Db 349 GSNQYLTSDNHQSPCAI-PEFDVTPPIDPGEVKNMELAEIDTMIPLNLESTKENTMDM 407
Qy 67 -EFTISPNNTPGDILFDLQGLPHLNPFLSH-----LSQMYNGWVGNMVRILLAGNAFSA 120
Db 408 YRVLTSDSADLSQPLICLSLSPAFDPRLSHTMLGVLNYYTHWAGSLKFTFLFCGSMAT 467
Qy 121 GKIIIVCCVPPGFTSSSLTIAQATLFPFHVIAVRTLEPIEMPLEDVNRVLYH--TNDNQPT 178
Db 468 GKILVAYAPPG-ADPTSRKEAMLGTHVINDLGQSSCTMVVPMWISNTYRQTQDSFTE 526
Qy 179 MRLVCMLY-----TPLRRTGGSGNSDSFVAVAGVLTAPSSDSSELSFLVPPTIEQKRAFT 233
Db 527 GGYISNFTQTRIVVPLST-----PKSMSMLGFV--SACNDFSVALLRDTT-----569
Qy 234 VPNIPLQTLNSRFPFSLIQGMILSPDASQVVOFQNGRCLIDGQLLCTTPATSGQLFRVRG 293
Db 570 -----HISQSALPQGIEDL-----TS 585
Qy 294 KINQARTNL-TEVDGKFPFAFDSAPVGPDPFGKCDHMR--ISKTPNNTSSGDPMR- 349
Db 586 EVAQGALTLSLPKQDLSLPDTKASGPA-----HSKEVPALTAVETGATNPLAP 633
Qy 350 SVSVQT-----NVQGFVPHLGSIQDFVFNHPTGDIYIGTIEWISQSPSTPPGTD 397
Db 634 SDTVQTRHVQRRSRSESTIESFFARGACVALLIEDN-----EQPTTRAKQL 680
Qy 398 INLWEIPDYGSSLSQAANLAPPVFPFGGEALVYFVSA-FPGPNRNSAPNDVPCLLPQBY 456
Db 681 FAWMRI-TYKDTVQLRRKLEFFTY-SRFDMEFTFVVNTANNGHALNQV-----Y 731
Qy 457 ITHFVSEQAPTMDAALLHVDP-----DTNRLGFEFLYPGGYLTCPVNGVGCAGPQ 508
Db 732 QIMYIPPGAPT-----PKSWDDYTWTQTSNPSIFTY-----GAAPA 768
Qy 509 QLPNGVFLFVSWVSRYQ-LKPVGTASTARSRLG 542
Db 769 RISVPVGLANAY-SHFYDGFPAKVLKTDANDQIG 802

RESULT 9
US-08-461-503-2
; Sequence 2, Application US/08461503
; Patent No. 5834302
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES
; TITLE OF INVENTION: FROM CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,503
; FILING DATE: 5-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-D-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-461-503-2

Query Match 4.8%; Score 140; DB 2; Length 2206;
Best Local Similarity 21.2%; Pred. No. 0.0006;
Matches 122; Conservative 62; Mismatches 221; Indels 170; Gaps 27;
QY 17 GAGQLVPEVNTADPLPMEPVAGPTTAVATAGV-NM-----IDPMVNNFVQSPQ----- 66
DB 349 GSNQYLTSNDHOSPCAI-PEFDVTPPIDIPGEVKNMELAEIDTMIPLNLESTKNTMDM 407
QY 67 -EFTISPNNTPGDILDLQGLPHLNPFLSH-----LSQMYNGWGNMVRILLAGNAFSA 120
DB 408 YRVTLSDSADLSQPIILCLSLSPAFDRLSHTMLGEVLYNTHWAGSLKFTFLFCGSMAT 467
QY 121 GKIIIVCCVPPGTSSTSLTIAQATLPHVIADVTLEPIEMPLEDVNRVLYH--TNDNQPT 178
DB 468 GKILVAYAPPG-AQPTTSKEMWGLTHVINDLGLOSSCTMNVVWISNVYRTQTQDSFTE 526
QY 179 MRLVCMCLY-----TPLRTGGSGNSDSFVAVGRVLTAPSSDFSFLVLPPTIEQKTRAPT 233
DB 527 GGVISMFYQTRIVVPLST-----PKSMMLGFV--SACNDFSRLLRDTT----- 569
QY 234 VNIPIQLTISNRFPGLIGMILSPDASQVQFQNGRCLIDGQLLGTTPATSGQLFRVNG 293
DB 570 -----HISQALPOGIEDL-----TS 585
QY 294 KINQARTLNL-TEVDGKPFMAFDSPAPVGFDFGKCDWHMR--ISKTPNNTSSGDPMR- 349
DB 586 EVAQGALTLSLPKQQSLPDTKASGPA-----HSKEVPALTAVETGATNPLAP 633

QY 350 SVSVQVT-----NVQGFVPHLGSIQPDEVFNHPTGDIYGTIEWISQSPSTPPGTD 397
DB 634 SDTVQTRHVQRRSRSESTIESFFARGACVAILIEVDN-----EQPTTRAOKL 680
QY 398 INLWEIPDYGSSLSQAANLAPVPFPFGGEALVYFVSA-FPGPNRSAPNDVPCLLPQBY 456
DB 681 FAWMRI-TYKTVQURRKLEFFTY-SRDMETTFVTANTFTNANNHALLNQV-----Y 731
QY 457 ITHFVSEQAPTMGDAALLHYVDP-----DTNRNLGEFKLYPGGYLTCVNGVGAGPO 508
DB 732 QIMYIPPGAPT-----PKSWDDYVTWQTSNPSIFITY-----GAAPA 768
QY 509 QLPNGVFLFVSWWSRFYQ-LKPVGCTASTARSLG 542
DB 769 RISVPYVGLANAY-SHFYDGFAPKVLKTDANDQIG 802
RESULT 10
US-08-465-250-2
; Sequence 2, Application US/08465250
; Patent No. 6136570
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release 1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,250
; FILING DATE: 6-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-E-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-465-250-2
Query Match 4.8%; Score 140; DB 3; Length 2206;
Best Local Similarity 21.2%; Pred. No. 0.0006;
Matches 122; Conservative 62; Mismatches 221; Indels 170; Gaps 27;
QY 17 GAGQLVPEVNTADPLPMEPVAGPTTAVATAGV-NM-----IDPMVNNFVQSPQ----- 66
DB 349 GSNQYLTSNDHOSPCAI-PEFDVTPPIDIPGEVKNMELAEIDTMIPLNLESTKNTMDM 407
QY 67 -EFTISPNNTPGDILDLQGLPHLNPFLSH-----LSQMYNGWGNMVRILLAGNAFSA 120
DB 408 YRVTLSDSADLSQPIILCLSLSPAFDRLSHTMLGEVLYNTHWAGSLKFTFLFCGSMAT 467
QY 121 GKIIIVCCVPPGTSSTSLTIAQATLPHVIADVTLEPIEMPLEDVNRVLYH--TNDNQPT 178

Db	468	GKILVAYAPPG-AQPPTSRKEAMLGTHVIWDLGLQSSCTMVVPMISNVYTRQTQDSFTE	526
Qy	179	MRLVCMVLY-----TPIRTGGSGNSDSFVVAGRVLTAPSSDFSFLVLPPTIEQKTRAF	233
Db	527	GGYISMFYQTRIVVPLST-----PKSMSLGFV--SACNDFSVLRLDIT	569
Qy	234	VENIPIQLTNSRFRPSLIQGMILSPDASQVQFQNGRCLIDQLLGTTPATSGQLFRVRG	293
Db	570	-----HISQALFQGIEDL-----	585
Qy	294	KINQOGARTNL--TEVDGKPFMAFDSPAPVPDFDKCDWHR--ISKTPNNTSSGDMR-	349
Db	586	EVAQGALTLSLPKQDSDLPDKASGPA-----HSKEVPALTAVETGATNPLAP	633
Qy	350	SVSVQV-----NVQGFVPHLGSIQDFEVENHPTGDYIGTIEWISOPSTPPGTD	397
Db	634	SDTVQTRHVQRRSRSESTIESFFARGACVALLIEDVN-----EQPTTRAQKL	680
Qy	398	INLWEIPDYGSSLSQAANLAPPVFPFGGEALVYFVSA-FPGNNRSAPNDVPCLLPQBY	456
Db	681	FAMWRI-TYKDTVQLRRKLEFFTY-SRFDMEFTFVVTANFTNANGHALNQV-----Y	731
Qy	457	ITHFVSEQAPTMGDALLHYVDP-----DTNRNLGEFKLYPGCYLTCVENVGVGAGPQ	508
Db	732	QIMYIPPGAPT-----PKSWDDYTWTQSSNPSIFYTY-----GAAPA	768
Qy	509	QLPLNGVLFVSVMSRFYQ-LKPVGTASTARSRLG	542
Db	769	RISVVPVGLANAY-SHFYDGFAXVPLKTDANDQIG	802
RESULT 11			
US-09-579-181-2			
; Sequence 2, Application US/09579181			
; Patent No. 6363372			
; GENERAL INFORMATION:			
; APPLICANT: Chrivia, John			
; APPLICANT: Yaciuk, Peter			
; TITLE OF INVENTION: SNF2 Related CBP Activator Protein (SRCAP)			
; FILE REFERENCE: 16153-4247			
; CURRENT APPLICATION NUMBER: US/09/579,181			
; PRIOR FILING DATE: 2000-05-25			
; PRIOR FILING DATE: 1999-05-27			
; NUMBER OF SEQ ID NOS: 17			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 2			
; LENGTH: 2972			
; TYPE: PRT			
; ORGANISM: Human			
US-09-579-181-2			
Query Match			
Best Local Similarity 4.2%; Score 123; DB 4; Length 2972;			
Matches 122; Conservative 57; Mismatches 206; Indels 192; Gaps 26;			
Qy	19	GQLVPEVNT----ADPLMEVAGP-TTAVATAGQVNMIDP--WIVNMFVQSPQGEFTIS	71
Db	809	GRTVVVNNPRAPLGFVDPVPPGPELSAQPTPGVPQVLPASLMVS---ASPAGPPLIP	865
Qy	72	PNNTPGDIILFDLQGLPHLNPFLSHLSOMYNGVGNMR-----VRIL	112
Db	866	ASRPPGPVLL-----PPIQPSGSLFQVLPGLVLSGTSRPPPTLSLKEPTTAPVRLS	920
Qy	113	LAGNAFSAAGKIIVCCVPPGF-----TSSSLITIAQATLFPHVIADVRLTPIEMPLEDV	165
Db	921	PAPPPGSSLLKPLIVPGYTFPPAAATTTTATAT-----TTAVPAPTAPQ-	970
Qy	166	RNVLYHTNDNQTM-----RLVCMLYTPLRTGGSGNSDSFVVAGRVLTAPSSDFS	216
Db	971	RLUIL--SPDMQARLPSGEVVISIGQLASLAQRVANAGGS-KPLTFQIQGNKLTLTGAQVR	1027
RESULT 12			
US-09-579-181-1			
; Sequence 1, Application US/09579181			
; Patent No. 6365372			
; GENERAL INFORMATION:			
; APPLICANT: Chrivia, John			
; APPLICANT: Yaciuk, Peter			
; TITLE OF INVENTION: SNF2 Related CBP Activator Protein (SRCAP)			
; FILE REFERENCE: 16153-4247			
; CURRENT APPLICATION NUMBER: US/09/579,181			
; CURRENT FILING DATE: 2000-05-25			
; PRIOR FILING DATE: 1999-05-27			
; NUMBER OF SEQ ID NOS: 17			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 1			
; LENGTH: 3118			
; TYPE: PRT			
; ORGANISM: Human			
US-09-579-181-1			
Query Match			
Best Local Similarity 4.2%; Score 123; DB 4; Length 3118;			
Matches 122; Conservative 57; Mismatches 206; Indels 192; Gaps 26;			
Qy	19	GQLVPEVNT----ADPLMEVAGP-TTAVATAGQVNMIDP--WIVNMFVQSPQGEFTIS	71
Db	955	GRTVVVNNPRAPLGFVDPVPPGPELSAQPTPGVPQVLPASLMVS---ASPAGPPLIP	1011
Qy	72	PNNTPGDIILFDLQGLPHLNPFLSHLSOMYNGVGNMR-----VRIL	112
Db	1012	ASRPPGPVLL-----PPIQPSGSLFQVLPGLVLSGTSRPPPTLSLKEPTTAPVRLS	1066
Qy	113	LAGNAFSAAGKIIVCCVPPGF-----TSSSLITIAQATLFPHVIADVRLTPIEMPLEDV	165
Db	1067	PAPPPGSSLLKPLIVPGYTFPPAAATTTTATAT-----TTAVPAPTAPQ-	1116
Qy	166	RNVLYHTNDNQTM-----RLVCMLYTPLRTGGSGNSDSFVVAGRVLTAPSSDFS	216
Db	1117	RLUIL--SPDMQARLPSGEVVISIGQLASLAQRVANAGGS-KPLTFQIQGNKLTLTGAQVR	1173
Qy	217	FLFL-----VPPTIEQKTRAFV-----PNIPIQLTNSRFR	247
Db	1174	QLAVGQPRFLQMPPTMNNNGIVVQVQAPRDGLTPVPPAPAPRPPSSGLPAVLNPR-	1232
Qy	248	PSLIQGMILSPDASQVQFQNGRCLIDGOLLGT--TPATSGQLFRVRGKINQARTNLNT	305
Db	1233	PTLTGRLPTP-----TLGTARAPMPTPTLVRPLLLKLVHSPSPPEVSA	1274

Wed Jun 2 09:13:27 2004

us-09-926-799-3.ra1

Search completed: June 1, 2004, 13:57:56
Job time : 15.0087 secs

```
Db      2599 NVGTGDHV 2606
| |||:
| |||:
| |||:

RESULT 15
US-09-976-594-15
; Sequence 15, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 15
; LENGTH: 2322
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6673549 1720920CD1
US-09-976-594-15

Query Match      3.88; Score 109; DB 4; Length 2322;
Best Local Similarity 22.08; Pred. No. 0.79;
Matches 131; Conservative 69; Mismatches 193; Indels 202; Gaps 35;

Qy      10 QSADGASGAG-----LVPEYNTADPLPMEFVAGPTT-----AVATAGQVNMIDPWIV 57
Db      1805 QGPAGASVAGPQTEAFITVRDNERPPQASVPLRLTRGSRAPISRAQLSVVD---- 1860
Qy      58 NNFVQSPQGEFTTSPNPTPGDILFDLQGLHNPFLSHLSQMYNGWGNM-----R 108
Db      1861 -----PDSAPGEIEYEVORAPH-NGFLS-----LVGGGLGPVTRFTQADVDS 1901
Qy      109 VRILLAGNAPSAGKIIIVCCVPPGPTSSSLTIAQAT-LFPHVIADVRLTEPIEMPLEDVNRN 167
Db      1902 GRLEFVANGSSVAGIFQLSKSDG-ASPLPMSLAVDILPSAI-EVQLRAPLEVP-----Q 1954
Qy      168 VLYHTNDNQPMRLVCMLYTP-----LRTGGGSGNSDFVVGAVLTAPSSDSFSLFLVP 222
Db      1955 ALGRSSLSQQQLRVVSDREPEEAAYRLIQGPYGH---LLVGGR-----PTSASFQF---- 2003
Qy      223 PTIEQKTEAFTVPNIPLQTLNSRFPSSLIOGMILSPDASQVQF-----Q 267
Db      2004 -QIDQGEVVFAFTNFS-SSHDFRVLALARGV----NASAVNVTVRALLHVWAGGPWQ 2057
Qy      268 NGRCLIDGQLL-----GTTP-----ATGQLFRV-RGKINQAGTL--NLTEVD 308
Db      2058 GATLRLDPTVLDAELANRTGSPFRLLGPRHGRVVRVPRARTEPGGSQLVEQFTQQD 2117
Qy      309 -----GKPFMAFSPAPVGPFDGCKDWMHRISKTPNTSSGDPMRSSVSQVQNVQ 358
Db      2118 LEDGRGLGLEVGRP--EGRAPGPAG-----DSLLELWQA 2149
Qy      359 GFVPHLGSIQFDEVFNHPTGDYIGTIEWISOP-----STPEGTDINLWEIPD- 405
Db      2150 GVPPAVASLDFAPEYNAAAPY--SVALLSVPEAAATEAGKPESSTPTGEPGPMASSPEP 2207
Qy      406 ---YGSLS-QAANTAPPVFPFGFGEALVYFSAFPGP-----NNESAPNDVPCLLPQ 454
Db      2208 AVAKGGFLSFLKANMFSVIIP-----MCLVLLALLILPLLFLYLRKRKNTKGHDVQVLTAK 2263
Qy      455 EVITHFVSQAPTMGDAALLHVVDPTNRLNGELKLYPGGY--LTCVPNGVGAGP 507
Db      2264 -----PRNGLAG-----DTETFR-----KVEFGQAIPLTAVP---GQGP 2294
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 34.9037 Seconds
(without alignments)
4368.312 Million cell updates/sec

Title: US-09-926-799-3

Perfect score: 2895

Sequence: 1 MMASKDAPQSDAGSAGAGQ.....QLKPGVTASTARSLGVRR 546

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2034.5	70.3	530	14	US-10-314-739-3
2	290	10.0	669	14	US-10-209-507-2
3	272	9.4	668	14	US-10-209-507-4
4	268.5	9.3	671	12	US-10-670-695-2
5	268	9.3	547	12	US-10-670-695-4
6	132.5	4.6	1194	12	US-10-282-122A-46163
7	127.5	4.4	2126	15	US-10-052-648A-39
8	123.5	4.3	1168	15	US-10-369-493-16733
9	123	4.2	2971	14	US-10-146-473-50
10	122.5	4.2	6310	12	US-10-282-122A-67793
11	115.5	4.0	1765	12	US-10-282-122A-48055
12	114.5	4.0	1194	12	US-10-282-122A-46577
13	113.5	3.9	1147	15	US-10-327-481A-38
14	111	3.8	1046	14	US-10-224-999A-3480
15	110.5	3.8	1394	12	US-10-381-247B-18

15	110.5	3.8	1435	12	US-10-276-774-2178	Sequence 2178, Ap
17	110	3.8	896	14	US-10-285-342-27	Sequence 27, Appl
18	109.5	3.8	2468	12	US-10-282-122A-66335	Sequence 66335, A
19	109.5	3.8	2468	14	US-10-246-330-4	Sequence 4, Appl
20	109	3.8	2322	10	US-09-919-039-15	Sequence 15, Appl
21	108.5	3.7	1025	11	US-09-834-309-5	Sequence 5, Appl
22	108.5	3.7	1106	14	US-10-157-031-30	Sequence 30, Appl
23	108	3.7	816	12	US-10-425-114-55912	Sequence 55912, A
24	107	3.7	1048	14	US-10-223-538-10	Sequence 10, Appl
25	106.5	3.7	3399	15	US-10-080-334-196	Sequence 196, App
26	106	3.7	1426	12	US-10-664-859-15	Sequence 15, Appl
27	106	3.7	1426	12	US-09-915-543-15	Sequence 15, Appl
28	106	3.7	1426	14	US-10-322-579-15	Sequence 15, Appl
29	105	3.6	575	9	US-09-738-626-4263	Sequence 4263, Ap
30	104.5	3.6	2514	12	US-10-282-122A-66121	Sequence 66121, A
31	104.5	3.6	2514	15	US-10-320-800-40	Sequence 40, Appl
32	102.5	3.5	2135	12	US-10-362-892-9	Sequence 9, Appl
33	102.5	3.5	2135	15	US-10-288-798-9	Sequence 9, Appl
34	102.5	3.5	2382	12	US-10-336-472-230	Sequence 230, App
35	102.5	3.5	2382	14	US-10-196-935A-20	Sequence 2, Appl
36	102.5	3.5	2382	15	US-10-052-648A-40	Sequence 40, Appl
37	102.5	3.5	3354	15	US-10-080-334-197	Sequence 197, App
38	102	3.5	944	14	US-10-174-677-101	Sequence 101, App
39	102	3.5	1185	15	US-10-259-194A-246	Sequence 246, App
40	102	3.5	2358	12	US-10-282-122A-45763	Sequence 45763, A
41	102	3.5	3930	12	US-10-282-122A-46817	Sequence 46817, A
42	101.5	3.5	680	16	US-10-389-566-429	Sequence 429, App
43	101.5	3.5	752	12	US-10-425-114-71415	Sequence 71415, A
44	101.5	3.5	1491	12	US-10-282-122A-65570	Sequence 65570, A
45	101.5	3.5	3013	12	US-10-282-122A-78257	Sequence 78257, A

ALIGNMENTS

RESULT 1

US-10-314-739-3
; Sequence 3, Application US/10314739
; Publication No. US20030129588A1
; GENERAL INFORMATION:
; APPLICANT: Estes, Mary K
; Jiang, Xi
; Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; Characterize No. US20030129588A1walk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Ave., N.W.
; CITY: Washington, D.C.
; STATE: <Unknown>
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/10/314,739
; APPLICATION NUMBER: US/10/314,739
; FILING DATE: 09-Dec-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,049
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311.023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX: <Unknown>

```
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 530 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-314-739-3

Query Match      70.3%; Score 2034.5; DB 14; Length 530;
Best Local Similarity 69.4%; Pred. No. 7.2e-183; Indels 17; Gaps 6;
Matches 379; Conservative 69; Mismatches 81;

QY 1 MMWASKDAPOSADGASGAGQLVPEVNTADPLPMEPVAGTTAVATAGQVNMIDPVIWVNF 60
Db 1 MMWASKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSTAVATAGQVNPIDPWIINF 60
QY 61 VQSPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLSOMYNGWGNMRVRIILAGNAFSA 120
Db 61 VQAPQGEFTISPNTPGDVLFDLSGLPHLNPFLSHLSOMYNGWGNMRVRIILAGNAFTA 120
QY 121 GKIIICVCPGFTSSSITIAQATLFPFHVADVRLTLEPIEMPLEDVNRNLYHTND-NQPTM 179
Db 121 GKIIICVCPGFTSSSITIAQATLFPFHVADVRLTLEPIEMPLEDVNRNLYHTND-NQPTM 180
QY 180 RLVCMLYTPRLTGGSGNSDSFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVNPIL 239
Db 181 RLVCMLYTPRLTGGGTG--DSFVAGRVMTCPSPDFNLFVLPPTVEQKTRFPLNPL 238
QY 240 QTLNSRFPPLSQMILSPASQVQFQNGRCLIDGQLLTTPATSGQLFRVRGKINQGA 299
Db 239 SLSLSNRAPLPISSMGISPNVQSVQFQNGRCLIDGRLVGTTPVLSHAKIRGTSN--G 296
QY 300 RTLNLTEDGKPFMAPSPAPVGPDPFGKCDMHMRISKTPNNTSSGDPMRSVSVQTNVQG 359
Db 297 TVINLTEDGTPFHPFEGPAPIGPDDLGGCDMHINMTQFGHSSQT-----QYDVTTPDT 351
QY 360 FVPHLSIQFDEVFNHPTGDIYGTIEWISQSPTPPGTDINLWIPDYGSSLSAANLAPP 419
Db 352 FVPHLSIQANGI---GSGNVGVLSWISPPSPSGQVDLWKIPNYSSTITEATHLAPS 408
QY 420 VFPFGCEALVYFVSAPFPGNRSAPNDVPCLLPQEIYTHFVSEQAPTMGDAALLHYVDP 479
Db 409 VYPPFGGEVLVFFMSKMPGP-----GAYNLPCLLPQEIYSHLASEQAPTVGEAALLHYVDP 464
QY 480 DTNRNLGEFKLYPGGYLTCVPNGVAGPQQPLNGVFLFVSVWSRFFYQLKPVGTASTARS 539
Db 465 DTGRNLGEFKAYPDGFLTCVPNGASSGPPQLPQINGVFVSVWSRFFYQLKPVGTASSARG 524
QY 540 RLGVRR 545
Db 525 RLGLRR 530
```

```
RESULT 2
US-10-209-507-2
; Sequence 2, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V
; FILE REFERENCE: 454313-3151.2
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
```

```
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
; US-10-209-507-2

Query Match      10.0%; Score 290; DB 14; Length 669;
Best Local Similarity 25.5%; Pred. No. 5.3e-18; Indels 130; Gaps 29;
Matches 153; Conservative 71; Mismatches 247;

QY 12 ADGAGAGQQLVPEVNT-ADPLPMEPVAGTTAVATAGQVNMIDPVIWVNFVQSPQGEFTI 70
Db 125 ADGDSI--TTPBQGLVGVIAEPSPQAATAAATAGKSDVSEW--ESFF-SFHTSVNM 179
QY 71 SPNTPGDIILFDLQGLPHLNPFLSHLSOMYNGWGNMRVRIILAGNAFSAKIIIVCCVPP 130
Db 180 STSETQKILFKQSLGFLNPFLEHLSKLYVAMSGSVDFRFSISGSGVGGKLAIVVVP 239
QY 131 GFTS-SSLTIAQATLFPFHVADVRLTLEPIEMPLEDVNRNLYHTNDNQPTMRVLVCMLYTPL 189
Db 240 GVDPVQSTMLQ--YPHVLFDARQVEPVIFSPDLRSTLYHLSMDTDTTSLVIMVYNDL 296
QY 190 -RTGGSGNSDSFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVNPILQTLNS-- 245
Db 297 INPYANDSNSGGCIVT--VETKPGDPFKHLLKPP-----GSMLTGCSIPSDILPKSSSL 349
QY 246 ----RFPSLIQMILSPASQVQFQNGRCLIDGQLLTTPATSGQLFRVRGKINQGA 301
Db 350 WIGNRYWSDITDFVIRP-----FVFQNRHFDNQ---ETAGWSTPRFR-----PIT 393
QY 302 LNLTEVDGKPP--MAFDSAPVGPDPFGKCDMHMRISKTP-----NNTSSGDP 347
Db 394 ITISESGSKLGTGATVDIYV-GIPD-GWPDITIGBELTPAGDYISITNGSGNDIATANA 451
QY 348 MRSVSQTNVQGVFPHLSIQFDEVFNHPTGDIY-GTIE--W-----ISQP 390
Db 452 YDSADVTNTTNF-----RGMYICGALQRAWGDKKISSATFITAIEG 495
QY 391 ST-PEGTDINLWELPDY-----GSSLSQAANLAPPVFPFGFEAL-----VYFVSAFP 437
Db 496 NTLKPSNTIDMTKLVQVQTHVGRDVQTSDDTLAILGVGTIGEQAIGNSRDSVVRISMLP 555
QY 438 GPNRNSA-----PNVPCLLPQEIYTHFVSEQAPTMGDAALLHYVDP-----DT 481
Db 556 ETGARGGNHPIFYKNSIKLGYVLSIDVPNSQILHTRQJSLNHYLLPPDSFAVYRIDS 615
QY 482 NRNLGEFKLYPGGYLTCVPNGVAGPQ-QLPLNGVFLFVSVWSRFFYQLKPVGTASTARS 540
Db 616 NGSWFDVGIDSDGFSFV---GVSSIPKLEFPLSAYMGI-----QLAKIRLASNIRST 665
QY 541 L 541
Db 666 M 666
```

```
RESULT 3
US-10-209-507-4
; Sequence 4, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V
; FILE REFERENCE: 454313-3151.2
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
```

```
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-4

Query Match          9.4%; Score 272; DB 14; Length 668;
Best Local Similarity 28.7%; Pred. No. 2.7e-16;
Matches 99; Conservative 39; Mismatches 135; Indels 72; Gaps 15;

QY 21 LVEPVNTA-DP-LPMEPVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEFTI-----70
Db 106 LIGEVAKADPNLPLFRLEADGGSVTPRQGT-----LVGGVIAEPNAQMSAVADVATGK 160
QY 71 -----SPNNTPGDILFDLQGLPHLNPFLSHLSOMYNGWGNMRVRL 112
Db 161 SVDSWEAPFSPHTSVNNSTSTQKILFKQSLGPLNPLYLTHLAKLYVAMSGSIEVRPS 220
QY 113 LAGNAFSAKIIIVCCVPPGFTS-SSLTIAQATLFPHVIAADVRLTEPIEMPLEDVNRVLYH 171
Db 221 ISGSGVFGKLAIIIVPGIDFVQSTSMQ---YHVLFDARQVEPVIFITDLNLSLYH 277
QY 172 TNDNQPTMRLVCMYTP-LRTGGGSGNSDSFVAGRVLTAPSSDFSFLVLPPTIEQKTR 230
Db 278 LMSDTHTTSLVMIYNDLINPYANDSNSSGCIVT--VETKPGDPFKFHLKPP-----GS 330
QY 231 AFTVNIPIQLTNS-----RPSLIQGMILSPDASQVVOFQNGRCLIDQLGTTTPA 283
Db 331 MTHGSIISDLIPKSSSLWGNRHSIDTDFVVKP-----FVFOANRHDFNQ---ETAG 382
QY 284 TSGQLFRVRGKINQKARTLNLTVEVDGKPF---MAFDSAPAPVGPDP 325
Db 383 WSTPRFR-----PIITVSEKSGKLGIGVATDSIVP-GIPD 418

RESULT 4
US-10-670-695-2
; Sequence 2, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; PRIOR FILING DATE: 2003-09-25
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 671
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-670-695-2

Query Match          9.3%; Score 268.5; DB 12; Length 671;
Best Local Similarity 26.8%; Pred. No. 5.7e-16;
Matches 110; Conservative 47; Mismatches 149; Indels 105; Gaps 18;

QY 10 QSADGASGAGQLVPEVNT-ADPLPMEPVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEF 68
Db 124 EADDGSITA---PEQGTWVGVIAPSAQMSAQMSTADMATGKSVDSSEW-----EAPF 170
QY 69 TI-----SPNNTPGDILFDLQGLPHLNPFLSHLSOMYNGWGNMRVRLILAGNAFSAK 122
Db 171 SPHTSVNNSTSTQKILFKQSLGPLNPLYLTHLAKLYVAMSGSIEVRPSISGSGVFGK 230

; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-4

Query Match          9.4%; Score 272; DB 14; Length 668;
Best Local Similarity 28.7%; Pred. No. 2.7e-16;
Matches 99; Conservative 39; Mismatches 135; Indels 72; Gaps 15;

QY 21 LVEPVNTA-DP-LPMEPVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEFTI-----70
Db 106 LIGEVAKADPNLPLFRLEADGGSVTPRQGT-----LVGGVIAEPNAQMSAVADVATGK 160
QY 71 -----SPNNTPGDILFDLQGLPHLNPFLSHLSOMYNGWGNMRVRL 112
Db 161 SVDSWEAPFSPHTSVNNSTSTQKILFKQSLGPLNPLYLTHLAKLYVAMSGSIEVRPS 220
QY 113 LAGNAFSAKIIIVCCVPPGFTS-SSLTIAQATLFPHVIAADVRLTEPIEMPLEDVNRVLYH 171
Db 221 ISGSGVFGKLAIIIVPGIDFVQSTSMQ---YHVLFDARQVEPVIFITDLNLSLYH 277
QY 172 TNDNQPTMRLVCMYTP-LRTGGGSGNSDSFVAGRVLTAPSSDFSFLVLPPTIEQKTR 230
Db 278 LMSDTHTTSLVMIYNDLINPYANDSNSSGCIVT--VETKPGDPFKFHLKPP-----GS 330
QY 231 AFTVNIPIQLTNS-----RPSLIQGMILSPDASQVVOFQNGRCLIDQLGTTTPA 283
Db 331 MTHGSIISDLIPKSSSLWGNRHSIDTDFVVKP-----FVFOANRHDFNQ---ETAG 382
QY 284 TSGQLFRVRGKINQKARTLNLTVEVDGKPF---MAFDSAPAPVGPDP 325
Db 383 WSTPRFR-----PIITVSEKSGKLGIGVATDSIVP-GIPD 418

RESULT 4
US-10-670-695-2
; Sequence 2, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; PRIOR FILING DATE: 2003-09-25
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 671
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-670-695-2

Query Match          9.3%; Score 268; DB 12; Length 547;
Best Local Similarity 26.9%; Pred. No. 4.7e-16;
Matches 104; Conservative 43; Mismatches 139; Indels 100; Gaps 16;

QY 34 EPVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEFTI-----SPNNTPGDILFDLQGLP 87
Db 21 EPSAQMSTADMATGKSVDSSEW-----EAFSFTSVNNWSTSETQKILFKQSLP 71
QY 88 HLNPLSHLSOMYNGWGNMRVRLILAGNAFSAKIIIVCCVPPGFTS-SSLTIAQATLPP 146
Db 72 LLNPYLEHLAKLYVAMSGSIEVRFSISGSGVFGKLAIIIVVPPGVDPVQSTSMQ---YP 128
QY 147 HVIADVRLTEPIEMPLEDVNRVLYHTNDNQPTMRLVCMYTP-LRTGGGSGNSDSFVAG 205
Db 129 HVLFDARQVEPVIFCLPDLRSTLYHLMSDTHTTSLVIMVYNDLINPYANDANSNCIVT- 187
QY 206 RVLTAAPSSDFSFLVLP-----TTEQKTRAFVNPINPLQTLNSRFPSLIQM 254
Db 188 -VETKPGDPFKFHLKPPGSMTHGSIIPSDILPKTSSLIWGN-----RYWSDITDF 237
QY 255 ILSPDASQVVOFQNGRCLIDQLGTTTPATSGQLFRVRGKINQKARTLNLTVEVDGKPF-- 312
Db 238 VIRP-----FVFOANRHDFNQ---ETAGWSTPRFR-----PISVTITEQNGAKLGI 281
QY 313 -MAFDSAPAPVGPDPFKGKCDWHMRISKTNNNTSSGDPMSVSVQTNVQGFVPHLSIQPDE 371
Db 282 GVATDYIYP-GIPD-----GW-----PDTTIPGELI----- 306
```

Db 372 VFNHPTGDIYGTIEWISQSPSTPGTD 397
QY 307 ----PAGDYAITNGTGNDDITATGYD 328

RESULT 6
US-10-282-122A-46163
; Sequence 46163, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46163
; LENGTH: 1194
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (3)..(3)
; OTHER INFORMATION: X=any amino acid
US-10-282-122A-46163

Query Match 4.6%; Score 132.5; DB 12; Length 1194;
Best Local Similarity 21.0%; Pred. No. 0.0097;
Matches 126; Conservative 60; Mismatches 203; Indels 211; Gaps 30;

QY 19 GOLVPEVN---TADPLMEFVAGPTTAVATAGQVN-MIDPWIVN-NFVQSPQ---GB-- 67
Db 242 GTITNEANVTYSQNPTEP---PTTTTTPPTNTSVRTALVNPFTKTVSQVADIGDII 298
QY 68 -FTISFNNT-----PGDILFDLQGLPHLPFLSHLSQMYNGWVG 105
Db 299 TTYTILPTNGISATNVIWTDIPAGTTIPNSVINGVQNPINP----- 344
QY 106 NMRVIRILLAGNAFSAKIIIVCCVPFGTSSSLTIAQATLFPH--VIADVR-----TLEPIE 159

Db 345 -----AGGIQVGTINAGSTTTTTFQVQVTSLPQNGVIRNIGNTTTFYQD-- 388
QY 160 MPLEDRVNLVYHTNDNQ-PMRLVCMLYPLRTGGSG-----NSDSFVVAG 205
Db 389 ---DPTKPTITTTNPTPTTVPINTAIINPIKADKTAVIDGDIITYITIFNNDGTVPAT 445
QY 206 RVLTAPSDFSFLFL-----VP---PTI-----EOKTRAFVTVPNIPLQTL 242
Db 446 NVIFTDSIPAGTTFIPNSVVLNNPVPNSPALGITVGLNPGETKTLISFQV----- 497
QY 243 SNRPFSLIQGMILSPDASQVQFQNGRCLIDGQLLGTTPATSGQLFVRGKINOGARTL 302
Db 498 ---RVTOIPAGGTITNEASTTYTQPDPTLPVPTTTEPTPTTS-----VTVNTATV 545
QY 303 NLTEVDGKPFMAFDSAPVGFDPFGKCDWHMRISKTPNNTSSGDPMRSVSVQTNVQGFVP 362
Db 546 NPTK-----SADRAFADIGDIITY-----TISLQNN--GTVP 575
QY 363 HLGSIQFDEVNHTPDYIG---TIEWISQSPSTPGTDINLM-----E 402
Db 576 ATNILLTDPIPNGT--FIPNSVTINGISQPNTPNSTGITVGLDPTPEAATISFQVQVIS 633
QY 403 IPDYGSSLQA-----ANLAPPVFPFGFGEALVYFVSAPFPNNRSAPNDV---PCILLQOE 455
Db 634 VPPHGLVENQGTVSFTHIVNENEP-----VTKTSPTKTTETAVNTIISTP----- 679
QY 456 YITHFVSEQAPTMDGAALLHYVDPTNRNLGEPKLYPGGYLTCVPGVAGPQOLPLNGV 515
Db 680 --TKTADKQLADIGDT--ITYT--ITFRNGTVPATNVTLIDSTPSGTTFIPDSVTINGV 733

RESULT 7
US-10-052-648A-39
; Sequence 39, Application US/10052648A
; Publication No. US20040005558A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steven
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gunther, Erik
; APPLICANT: Kekuda, Ramesh
; APPLICANT: MacDougall, John R.
; APPLICANT: Mehraban, Fuad
; APPLICANT: Patturajan, Meera
; APPLICANT: Rothenberg, Mark
; APPLICANT: Shinkets, Richard
; APPLICANT: Smithson, Glennnda
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Stone, David J.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerkusen, Bryan D.
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; FILE REFERENCE: 21402-250 (CURA-550)
; CURRENT APPLICATION NUMBER: US/10/052,648A
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,920
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/284,549
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/303,229
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,605
; PRIOR FILING DATE: 2001-01-23

; PRIOR APPLICATION NUMBER: 60/269,098
 ; PRIOR FILING DATE: 2001-02-15
 ; PRIOR APPLICATION NUMBER: 60/264,159
 ; PRIOR FILING DATE: 2001-01-25
 ; PRIOR APPLICATION NUMBER: 60/265,517
 ; PRIOR FILING DATE: 2001-01-31
 ; PRIOR APPLICATION NUMBER: 60/271,855
 ; PRIOR FILING DATE: 2001-02-27
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 97
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 39
 ; LENGTH: 2126
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 US-10-052-648A-39

Query Match 4.4%; Score 127.5; DB 15; Length 2126;
 Best Local Similarity 20.9%; Pred. No. 0.07;
 Matches 135; Conservative 63; Mismatches 246; Indels 201; Gaps 30;

QY 3 MASKAPQADGASGAGQLVPEVNTADP-----LPME---PVAGPTTA 42
 Db 937 MLESDVSEPEGDQGLSLOKDDYGFQSKLEGFEKQPIAVSSMPQOIGVPTSLTQV 996
 QY 43 VATAGQVMIDWIVNFWQSQGEFTISPNTPGDILFDLQGLHNLPHLSQMYNG 102
 Db 997 VHSAGR---RFIVSPVESRLRESKIFTSEIPDPVAASTSQPGMN--LSHSASS---1046
 QY 103 WVGNNRVRILLAGNAFSA---GKII--VCCVPPGFTSSSLTIAQATLPHPHVADVRLTE- 156
 Db 1047 -----LSLQQAFSELKHQGTGPNAPFNHNPPTFSP---FLTSIAGVQTAA 1094
 QY 157 -----PI-EMPLEDVRNLVHTNDNQPTMRLVCMLYPLRLTGGSGNSDSFWAGRYL 208
 Db 1095 STFSVSVPITSSPLNDISVWQSEGALEPTDKGI-----CGVTSTGVASGGLT 1144
 QY 209 TAPSDPSFLVVPPTIEOKTRAFVNPILQTLNSRPPSLIQGMILSPDASQVVQFQN 268
 Db 1145 TILSVSE-----TPTLSSAVSSSTAPAV--VTVSTTSQP--VQAPTSGSIASSTGSFPS 1193
 QY 269 GRCLIDQLLGYTPATSGQLFVRGKINGARTLNLTEDVCK-----PF 312
 Db 1194 -----GTFSTTGTTVSSVAVPNNAKPTVLLQQVAGNTAGVAIVTSTVSTTPFPA 1243
 QY 313 MAPDSAPVG-----PFDGK---CDWHMRIKSTNNNTSSG-----DPMRSVSQVOTNVQ 359
 Db 1244 MASQPSLPLGSSAPTAEVTVVSNH-SLDKASHSTAGLGLFCAPSSSSSGTAVSS 1302
 QY 360 FVPHLGSIQDFVFNHP--TGDYIGTIEWISQFSTPPGTDINLWEIPDYGSSLSQAANLA 417
 Db 1303 SVSQPGIV-----HPLVISAIASTPVLQPAVPTSTPL-----LPQVFNIP 1344
 QY 418 PVFPP-----PGGEALVYFVSAPFGNNRSAPNDVPCLLPQBYITHFVSEQAFTMGDA-- 471
 Db 1345 PLVQPVANVAVQOTLIH-----SQPQ--PALLPNOPHTCPMDADTOSKAPG 1391
 QY 472 -----ALLHYVDPD--TNRNLGEFKLYPGGYLT 497
 Db 1392 IDDIKTELEKRLSRFSEHSSGTQHSVSLETPLVVEVTGIPITTAAPSKLMTSTTST 1451
 QY 498 CYPNGVAGPQOPLNGVFLFVSWSRFYQLKPVGT-----ASTA 537
 Db 1452 CLP-----PTNLPLGTAGMPVMPVGTGQVSTPGTHASAPASTA 1490

RESULT 8

US-10-369-493-16733
 ; Sequence 16733, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.

RESULT 9

US-10-146-473-50
 ; Sequence 50, Application US/10146473
 ; Publication No. US20030108888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Scanlan, Matthew
 ; APPLICANT: Gout, Ivan

; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 16733
 ; LENGTH: 1168
 ; TYPE: PRT
 ; ORGANISM: Bacillus thuringiensis
 US-10-369-493-16733

Query Match 4.3%; Score 123.5; DB 15; Length 1168;
 Best Local Similarity 18.9%; Pred. No. 0.066;
 Matches 120; Conservative 96; Mismatches 249; Indels 169; Gaps 27;

QY 3 MASKAPQADGASGAGQLVPEVNTADPMPPEFVAGPTTAVATAGQVN--MIDPMI--- 56
 Db 91 LASSIPQ-----GAIINQALTSYTYIVDPSPQPPVTATSSNTVNTAVVDASLSVIK 142
 QY 57 -VNNFVQSPQGEFTI-----SPNTPGDILFDLQGLHNLPHLSQMYNGVGNMVR 110
 Db 143 STDLSLVQSDGTITVTVVQNNGTNTANTVTLTLVPE-----180
 QY 111 ILLAGNAFSAGKIIIVCCVPPGFTSSSLTIAQATLPHPHVADVRLTEPIEMP-----LE 163
 Db 181 -----GTAP-----IPNSVTISGVSPGAD--PNVGIPLNSITSPSIVTVFOVIV 225
 QY 164 DVNRNLVHTNDNQPTMRLVCMLYPLRLTGGSGN-----SDSFVAGVRLTAPSSDFSF 217
 Db 226 SIPSTVQPSNTARIDYTFIADPTAFIISRTIISNPAFTQISDATILSLKAVNAPQTTGD 285
 QY 218 LFLVPTTIEOKTRAFVNPILQTLNSRPPSLIQGMILSPDASQVVOFQNGRCLIDG-O 276
 Db 286 ILTYITITLE-----NTGNIPATNL-----IPSDTIPQDTPFVENSFTLNGTA 327
 QY 277 LGITTP-----ATSQLFRVRGK-----INQARTLNLTEDVGKPPMAFD 316
 Db 328 ILGANPNVGVTLPLAANATHLISQILIKOSFSRESITNQSNNTTYTIQDFGQPPITET 387
 QY 317 SPAPVGFDPFGKCDWHMRIKSTNNNTS--SGDPMRSVSQVTVQGVFVPHLGSIQDEVEN 374
 Db 388 STSNIVITNFVQA--QLTITKTSNPTTVDVGITILYISEVKNI-GNVDAINIVFTDSI-- 442
 QY 375 HPTG-----DYIGTIEWISQFSTPPGTDINLWEIPDYGSS--LSQAANLAPVPFPFGGE 427
 Db 443 -PAGTTFVPDSV-TINGVLQGVNPENGIPIGTIPANSSKTLIFQVQTNPPPTETEIVNQ 500
 QY 428 -ALVYFVSAPF--GPNNSAPNDVPCLLPQBYITHFVSEQAFTMGDAALLHYVD-PDTHR 483
 Db 501 SSATQYVSITPAPPNRSANSNIIVTSLQN--ANIIISVKSDITFASIGQIITVTNLQ 558
 QY 484 NLGEFKLYPGGYLTCVPNGV-----GAGPO-----508
 Db 559 NIGTVPANNTVFIDNIPGTGTFIEDSLAINNVIQGANPENGVTLGTQNETVTVISFOV 618
 QY 509 ---QLPLNGVFLFVSWSRFYQLKPVGTASTARS 539
 Db 619 QLTNIPGVNVINISDTSVEYQIDPSSQIIQRRS 652

```

; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tsung
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 50
; LENGTH: 2971
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-146-473-50

Query Match      4.2%; Score 123; DB 14; Length 2971;
Best Local Similarity 21.1%; Pred. No. 0.31;
Matches 122; Conservative 57; Mismatches 206; Indels 192; Gaps 26;

QY 19 GQLVPEVNT---ADPLMEPVAGP-TTAVATAGOVNMIDP--WIVNMFVQSGEFTIS 71
Db 809 GRTVVVNNPRAPLGPVVRPPPELPSAQTPGFPVQVLPASLMVS---ASPAGPLIP 865
QY 72 PNNTPGDILQLGPHLNPFLSHLSQMYNGWVGNMR-----VRIL 112
Db 866 ASRPPGPVLL-----PPLQNSGSLQVLPSPGLVSGTSRPTTSLKPTTPAPVRLS 920
QY 113 LAGNARSAGKIIVCVPPGP-----TSSSLTIAQATLPHVIADVRTLEPIEMPLEDV 165
Db 921 PAPPGPSLSLLKPLTVPPGYTFPPAAATTTTATAT-----TTAVAPTAPQ- 970
QY 166 RNVLVYHTNDNQTM-----RLVCMLYTPLRTGGSGNSDSFVVAGRVLTAPSSDFS 216
Db 971 RLIL--SPDQARLPSEGVVSIQGLASLAQRPNVANAGS- KPLTFQIQGNKLTLTGAQVR 1027
QY 217 FLFL-----VPTTIEQKTRAFV-----PNIPLQLNSRF 247
Db 1028 QLVAGQPRPLQMPMTMNTGVVXIVVRQAPRDLTPVPLAPAPRPPSGGLPAVLNPR- 1086
QY 248 PSLIQGMILSPDASOVQFQNGRCLIDQLLGT--TPATSGQLPRVRGKINQOARTLNT 305
Db 1087 PTLTPGRLTP-----TLGTARAMPPTLVRPLKLVHSPSPVSA 1128
QY 306 EVDGKPFMAFDSP--APVGFPDFGKCDWHMRISKTPNNTSGDPMRVSQVTVNQGVEPH 363
Db 1129 SAGAAPLTISPLHVPSSLP--GPASSPMPI---PNSSPLASPVSS--TVSVPLSSSLP- 1181
QY 364 LGSIQFDEVFNHPTGYITTEWISQSPSTPGTDLNLWEIPDYGSSLSQAA-----NL 416
Db 1182 -----ISVPTLPAPASAPLTIP-ISAPLTVSASGALLTSV 1217
QY 417 APPVPP-----PGFGEALVYFVSAPPG 438
Db 1218 TPPLAPVVPAPPGPSPLOPSGASPSASALTIGLATAPSLSSSQTPGHPHLLAPTSSHPVG 1277
QY 439 PNNRSAPNDVPCLLPQEIY-THFVSEQAPTWGDAAALL 474
Db 1278 LNSTVAPACSPVLVPSALASPPSAPNPAPAQASLL 1314

RESULT 10
US-10-282-122A-67793
; Sequence 67793, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Lianguo
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlssen, Kari

; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 67793
; LENGTH: 6310
; TYPE: PRT
; ORGANISM: Pseudomonas putida
; US-10-282-122A-67793

Query Match      4.2%; Score 122.5; DB 12; Length 6310;
Best Local Similarity 19.5%; Pred. No. 1.1;
Matches 106; Conservative 63; Mismatches 197; Indels 177; Gaps 21;

QY 15 ASGAQQLVPEVNTADPLMEPV-----AGPTTAVATAGOVNMIDPWVNNFVQSGPGE 67
Db 1776 ASSALQFEPVPTTAPASPSDLFAEDGSSISGCTAPAGTRVEVDHANGTLTGTVVAGPDGS 1835
QY 68 FTIS--PNNTPGDILFDLQGLHNPFLSHLSQMYNGWVGNMRVRIILAGNAFSAKGIIV 125
Db 1836 FTVTLDPDPAQTNGELL-----DVVAIDGGVSSL-----1863
QY 126 CCVPGFTSSSLTIAQATLPHVIADVRTLEPIEMPLEDVRNVLVYHTNDNQTMELVCM 185
Db 1864 ---PQITAPDITAPAA-----PTLVINADGSSVVTGABPGSTVRVLA--1904
QY 186 YTPLRTGGSGNSDSFVVAGRVLTAPSSDFSLFVLVPPPTIEQKTRAFVVPNIP--LOTLS 243
Db 1905 -----ADGTVLGSVIVGATGSFS--ITLDPQIDGVEVLQVLTATDAAGNASTAS 1951
QY 244 NSRFPSSLIOG-----MILSPDASQV-----VQFQNGRCLIDGQLLGTTPAT 284
Db 1952 NLTPADIDGDTTPPEAPTNLVINPAGSOLTCRGEAGTSVQVRDG-----AGTVVAT 2003
QY 285 SGQLFRVRGKINOG---ARTLNLTVEVDGKPFVA-----FDSAPVGFDPFGKCDWHM 333
Db 2004 -----GTVPDGTGTFAILNPAITDGTSTLQVTLTDAAGNVQSGSVATPD-----L 2048
QY 334 RISKTNN-----TSSGDDPMRVSQVTVNQGVEPHLGSIQFDEVFNHPTGDIYGT 384
Db 2049 LAPQPTFELALIDGVTFTGRGECATVQVR-DATGSLIGTLVNE-----GTF 2096
QY 385 EWISQSPSTPPGTDINLWEIPDYGSSLSQAANLAPPVFP-----GFGEA 428
```


Db 2097 SVTLFPAQANGALDIRVVDAGNSSTPLAFTAPDVTTPAAVSNVVVGADGLVLSGRGEA 2156
QY 429 LV-----YFSAFPCPNRRSAPNDVPCLLPOEYITHFVSEQAPTWGDAALLHYVD 478
Db 2157 GATVQVRDANGSVIGTATVGANGTFLVDLTPAAQFCEQLSLVQTDASGNASEA--LOYEI 2214
QY 479 PDT 481
Db 2215 PAT 2217

RESULT 11
US-10-282-122A-48055
; Sequence 48055, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48055
; LENGTH: 1765
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-282-122A-48055

Query Match 4.0%; Score 115.5; DB 12; Length 1765;
Best Local Similarity 21.0%; Pred. No. 0.72;
Matches 118; Conservative 61; Mismatches 203; Indels 179; Gaps 27;

QY 7 DAPQSDAGSAGQLVPEVNTADPLMEPVAGPTT--AVATAGQVNMIDPW---IVNPFV 61
Db 870 DVPSVAFTVGNAGAVSVIRS-----KAVGLVLTLPAAASVAVTVRLCAPSPKPVGVNV 923
QY 62 QSPQGEFTISPNTPGDILFDLQGLPHLPFLSHLSQMTNGVGNVRILLAGNAFSAG 121
Db 924 QPFDGSVAVVPPNV-----VPSYTLTVL----- 946

QY 122 KIIVCCVPFGFTSSSLTTIAQATLPHVVIADVRTLBP-----IEMPLEDRVNLVH-----TND 174
Db 947 -----PASAVPLNVGVASSVLPPEMIAP--TTEPTSSVTVMTCAPGAVVSPVTVTEA 997
QY 175 NQPTMRLVCMYLTPLRTGGSGNSDSFVVA--GRVLT-----APSSDFSLFLVP-----PT 224
Db 998 DGFVLPAASAVVATPITVPAGSGVEGVYVHAPLGSAVTVASGVSPFVSMLTVPASAVPD 1057
QY 225 IEQKTRAFTPVNIPLQTLTSLNSRFPSSLIQGMILSPDASQVVFQNGRCLIDGOLLG--TTP 282
Db 1058 SDVPSVAVTVGN-----AGAVVSVIYSSKAV-----LGTLTLP 1089
QY 283 ATSGQLFRVRGKINGKARTLNITEVDGKPFMAFDSPAPVGFDFGKCDHMRISKTPNN- 341
Db 1090 AVS-----VAVTVRLCAPSPKPVGVN-----VQFPDGSVAV-----VVPNV 1126
QY 342 ----TSSGDPMSRSVSQTNVQGFV--PHLGSIQFDEVFNHPTGDIYGTIEMISQBST-PP 394
Db 1127 VPSYTLTVLPASAVPLNVGVASSVLPPEMIA-----PTEPTSS--VTVPMTGAP 1174
QY 395 GTDINLWEIPDYGSSLSQAANLA--PPVPPPGFGEALVY-----FVS 434
Db 1175 GAVVSPVTVTEADGPVLPAAVAVTPTITVPAGSGVEGVYVHAPLGSAVTVASGVSPFVVS 1234
QY 435 APP-GPNNRSAPNDVPCLL-----POEYITHFVSEQA-----PTMGDAALLHYVDPTN 482
Db 1235 MLTVAPASAVPDNDVPSVAVTVGIAGAVVSVIRSKAVLGTLTLPVAVSVAVTVRLCAPSPS 1294
QY 483 RNLGSEKLYPGGYLTCVNGV 503
Db 1295 AVVGVNVQFPDGSVAVVPSNV 1315

RESULT 12
US-10-282-122A-46577
; Sequence 46577, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09

Wed Jun 2 09:13:27 2004

PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 46577
LENGTH: 1194
TYPE: PRT
ORGANISM: Bacillus anthracis
US-10-282-122A-46577

Query Match 4.0%; Score 114.5; DB 12; Length 1194;
Best Local Similarity 19.6%; Pred. No. 0.49;
Matches 87; Conservative 55; Mismatches 154; Indels 147; Gaps 21;
QY 20 QLVPEVNTADPLPMEBVGAGTTAVATAGQVNMIDPWVNFVQSPQGEFTTSPNN----- 74
DB 555 QNVGNINATDVITDTPAGTTFIPNS-----VTINGVSSPCA-----NPNSGVNVG 601
QY 75 --TPGDILFDLQGLPHLPFLSHLSQMYNGWGNMRVILLAGNAFSAAGKIIIVCCVPP-G 131
DB 602 TVTPGGI-----VTLT-----QVTVTALPPDG 624
QY 132 FTSSSTUIAQ-----ATLPHVIADVLTLEPIEMP-----LEDVRNVLXHTN 173
DB 625 IIKNTATVTVTFQNFCEPEITITDPTTVEVSUITPTPNPKLADKQIVDINEIITYTV 684
QY 174 DNPOTMRLVCLVTLPLRTGGSGNSDSFVAGRVLTAPSSDFSLFLVPPTI-----EQKT 229
DB 685 TFO-----NRGVPATSVIIVTDPLA--NGLTFVPGTVILNGIPDL 722
QY 230 RAFTVNPNIQTLNSRFPSLIOGMILSPDASQVQFONGRCLIDQGLLGTTPATSGQLF 289
DB 723 GANPVEGIPVGT-----VNPNDTIVQFQ-----ARVTSVPPGGII 758
QY 290 RVRKINGARTLNLTEVDGKPFMAFDSAPVGFPGKCDMHRISKTPNNTSSGDPMR 349
DB 759 R-----NOATVTFYEPGEPPVITDPTTINTDVTALLNPQKATPFTVLGDIIT 813
QY 350 -SVSQVQNVQGVPHLGIQDFVFNHTG-DYTG-----TIEWISQSPSTPGTDINLWEIP 404
DB 814 YTISLQNT--GTIPANNILVSDPI---PTGTSFIONSVTINNVSQPTANPETGI---QIP 865
QY 405 DYSSLSQAAN---LAPVFPFG 424
DB 866 TLPSPSESATISFHVLTSPSPSG 888

RESULT 13
US-10-327-481A-38
Sequence 38, Application US/10327481A
Publication No. US20040001864A1
GENERAL INFORMATION:
APPLICANT: King, Andrew M. Q.
APPLICANT: Burman, Alison J.
APPLICANT: Audonnet, Jean-Christophe F.
APPLICANT: Lombard, Michel F.A.
TITLE OF INVENTION: Vaccine against Foot-and-Mouth Disease
FILE REFERENCE: 454313-3178
CURRENT APPLICATION NUMBER: US/10/327,481A
CURRENT FILING DATE: 2002-12-20
PRIOR APPLICATION NUMBER: PCT/FR01/02042
PRIOR FILING DATE: 2001-06-27
PRIOR APPLICATION NUMBER: FR 0008437
PRIOR FILING DATE: 2000-06-29
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 38
LENGTH: 1147
TYPE: PRT
ORGANISM: Foot-and-mouth disease virus
US-10-327-481A-38

Query Match 3.9%; Score 113.5; DB 15; Length 1147;
Best Local Similarity 21.2%; Pred. No. 0.57;
Matches 109; Conservative 45; Mismatches 188; Indels 171; Gaps 23;
QY 36 VAGPTTAVATAGQVNMIDPWVNFVQSPQGEFTTSPNNTGDIILFDLQGLPHLPFLSH 95
DB 129 VAGPNTSGLETRVQ-----AERFFKFLFDWT---TDKPPGYLTKLELPTDHDHGVGH 179
QY 96 LSQMY-----NGWGNMRVILLAGNAFSAAGKIIIVCCVPPGFTSSSLTTIAQATFLPHVIAD 151
DB 180 LUDSYAMRNGW-----DVEVSAVGQFNGCCLLVAMVPEWKAFTREKYQLTLFPHQFTS 235
QY 152 VRT--LEPIEMPLEDVRNVLXHTNDQPTMRVLVCLMYTLPLRTGGSGNSDSFVAGRVLT 209
DB 236 PRINMTAHITVPYLVGNR--YDQYKHKPWTLLVWVLSPL----- 273
QY 210 APSDSFSLFLVLPPTTIEQKTRAFVTVENIPLQTLNSRFPSSL-----IQGMILS 257
DB 274 -----TVSNTAAPQIKVYANIAITYVHVHAGELPS 302
QY 258 PDASQVQFONGRCLIDGQLLGTTPATSGQLFRVRGI-----NOGARTNLNTEV----- 307
DB 303 KEGIFPVACADGY-----GGLVTTDPKTADEVY---GKVYNPPKTYNPGRFTNLLDVAEAC 355
QY 308 -----DGKPFM---AFDSPAPVFPDFGKCDMHRISKTPNNTSSGDMRMSVSVQTNV 357
DB 356 PTFLEFDDGKPYVVVTRADDTL-----LAKFQVSLAAKHSNTYLSG-----IA 399
QY 358 QGFVPHLGSQDFVFNHTGDIYGTIEWISQP---STPPGT-----DINL--- 400
DB 400 QYTYQISGTLNLHFMFTGSTDASKARYVAVIPQVETFPDTPBEEAAHCHAEWDTGLNSK 459
QY 401 --WEIP-----DYSSLSQAANLAPVFPFGFGEALVYFVSAPFGPN-----NRSAPNDV 448
DB 460 FTFSIPYVSAADYATASDTAETTNVQ-----GWVCYVQITHGKAENDTLLVSASAGKDF 514
QY 449 PCLLPQEVITHFVSEQAPTMGDAALLHVDPDT 481
DB 515 ELRLPIDPRT-----QTTTGESA-----DEVT 537

RESULT 14
US-10-224-999A-3480
Sequence 3480, Application US/10224999A
Publication No. US20030171318A1
GENERAL INFORMATION:
APPLICANT: Myriad Genetics, Inc.
APPLICANT: Morham, Scott
APPLICANT: Zavitz, Kenton
APPLICANT: Hobden, Adrian
TITLE OF INVENTION: Composition and Method for Treating Viral Infection
FILE REFERENCE: 5004.01
CURRENT APPLICATION NUMBER: US/10/224,999A
CURRENT FILING DATE: 2003-03-03
PRIOR APPLICATION NUMBER: US 60/313,695
PRIOR FILING DATE: 2001-08-20
NUMBER OF SEQ ID NOS: 3484
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3480
LENGTH: 1046
TYPE: PRT
ORGANISM: Human herpesvirus 5
US-10-224-999A-3480

Query Match 3.8%; Score 111; DB 14; Length 1046;
Best Local Similarity 23.2%; Pred. No. 0.85;
Matches 109; Conservative 50; Mismatches 181; Indels 130; Gaps 28;
QY 17 GAGOLPEVNNTA---DPLPMEPVAGPT-----TAVATAGQVNMIDPWVNFVQSPQGEF 68
DB 580 GAGVNVDPAGAGAAILTTFTFVNPSTAPAPPTPTTFACTQTPVNGNSPW-----APTAPL 632
QY 69 --TISPNNTPGDIILFDLQGLPHL--NPFLSHLSQMYNGWGNMRVILLAGNAFSAAGKII 124

Db 633 PGDMNPANPRERAWALK-NPHLAYNPP-----RMPTSTTSQNNVS----- 673
QY 125 VCCVPFGTSSSLITIAQATLPHVIAVRL--EPIEMPLEDVNRVLYHTNDNQPTMLRV 182
Db 674 ---TTPRRSTPRAAVTQTASQNADEVWALRDQTAESFVEDSEE-----EDDD----- 719
QY 183 CMLYPLRTGGSGNSDSFVAVGRVLTAPSSDFSLFLVPP--TIEQKTRATVP----- 235
Db 720 -----SSDTGSVVSLSGH--TTPSSDYNDV-ISPSSQTPEQST-----PSRIRK 759
QY 236 ---NIPLQTLNSRPSLIQCMILSPDAS--QVVFQNGRCLIDCOLLG---TTPATSG 286
Db 760 AKLSSPMITTSQKP-VLGGKRVATPHASARAQTVTSTPVQGRVEKQVSGTSTVPATLL 818
QY 287 QLFVRVGRK-----INQARTLNLTVDGKPFMAFDSP-----APVGPDPFGKCDWHR 334
Db 819 QQPASSKTTSSRNVTSRGARTSASARQPSASVLSPTEDDVVSPVTSF-----LSMLSS 874
QY 335 ISKTNTNTSSGDMRSVSQTVNVQGVPHLGSIQDFEVNHP-----TGDIYIGTIEMWISQ 389
Db 875 ASPSPAKSAPPSPVKGSRGVPSLKPTLGG---KAVVGRPPSPVSGSAPCRLSGTSR 931
QY 390 -----PSTPPGTDINLWEIPDYGSSLSQAANLAPPV-----PFPQGEAL 429
Db 932 AASTTPTYPAVTV-----YPPSSTAKSSVSN-APPVAPSILKPGASAAAL 976

RESULT 15

US-10-381-247B-18
; Sequence 18, Application US/10381247B
; Publication No. US20040073001A1
; GENERAL INFORMATION:
; APPLICANT: Akiyama, Tetsu
; APPLICANT: Adachi, Shungo
; APPLICANT: Kyowa Hako Kogyo Co., Ltd.
; TITLE OF INVENTION: beta-catenin Nuclear Localizing Protein
; FILE REFERENCE: 082394-000000US
; CURRENT APPLICATION NUMBER: US/10/381,247B
; CURRENT FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: JP 2000-287876
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: WO PCT/JP01/08140
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 1394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-381-247B-18

Query Match 3.8%; Score 110.5; DB 12; Length 1394;
Best Local Similarity 20.2%; Pred. No. 1.5;
Matches 97; Conservative 41; Mismatches 160; Indels 183; Gaps 25;
QY 11 SADGASGAGQLVPEVNTADPLEMEPVAGPTTAVATAGOVNMIDPMVNNFVOSPOGEFTI 70
Db 1049 SNNNMPGKG-----INTQNPRISSGP--NPVVPMTLSPMGMTQPLSHSNQMFSPN---AV 1098
QY 71 SPNNTPGDILFDLQGLPHL---NPLSHLSQMYNGWGVNMVRVRIILAGNAFSAAGKIIVCC 127
Db 1099 GENIIPPHGV---PMGPGMLSHNPIMGHGSQ----- 1125
QY 128 VPPGFTSSSLITIAQATLPHVIAVRLTLEPIEMPLEDVNRVLYHTN----- 173
Db 1126 BPVMPVQGRMGFPQG--FP-----FVOSPPQQV--PPPHNGPSGGQSPFGMGF 1171
QY 174 -----DNOPTMRLVCMLYTLRTGTGGSGNSDSFVAVGRVLTAPSSDFSFLVLRPT 224
Db 1172 PEGPLGRPSNLPOSSADAALCKP-----GGPGGDSFTVLGNSMPSVFTD-----PD 1219
QY 225 IEQKTR--AFTVENIPLQTLNSRFPFSLIQCMILSPDASQVVQ-FQNGRCLIDGQLLGT 281

Db 1220 LQEVIRPGATGIPFDLSRIIPSEKP-----SQTLOYPPRGVPGRKQDQGGPG 1267
QY 282 PATSQQLFRVRGKINQAGARTLNLTVDGKPFMAFDSPAPVGPDPFGKCDWHRISKTPNN 341
Db 1268 PGFS-----HMQGMGGEQAPRMCLA-----LPGMGGPGVGTPTDI-----PLG 1305
QY 342 TSSG-----DPMRSVS--VQTVNOGVFVPHLGSIQDFEVNHPITGDYIGTIEMWISOPSTPPT 396
Db 1306 TAPSPMGHNPMPAPPAPFLQGMGMG--PHRMMS-----PAQSTMPGQ 1344
QY 397 DINLWEIPDYGSSLSQAANLAP-----PVFPFGGALVYFVSAPGCPNNR 442
Db 1345 -----PTLMSNPAAVGMIFGKDRGPAGLYTHFGPVGSPG-----MMMSMQGMGMPNRT 1393
QY 443 S 443
Db 1394 S 1394

Search completed: June 1, 2004, 14:04:31
Job time : 36.9037 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:34:56 ; Search time 6.59597 Seconds
(without alignments)
4310.252 Million cell updates/sec

Title: US-09-926-799-3
Perfect score: 2895
Sequence: 1 MMASKADAPQADGASGAGQ.....QLKPVGTASTARSLGVRR 546

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2887	99.7	546	1 COAT_SOUV3	Q04542 southampton
2	1114	38.5	539	1 COAT_LORDV	P54635 lordsdale v
3	328	11.3	2344	1 POLN_RHDV	P27410 rabbit hemo
4	299	10.3	703	1 COAT_SMSVA	P36285 san miguel
5	293.5	10.1	702	1 COAT_SMSV1	P36284 san miguel
6	284	9.8	668	1 COAT_FCVF4	P27405 feline cali
7	269.5	9.3	2208	1 POLN_WANCV	Q69014 manchester
8	268.5	9.3	671	1 COAT_FCVF9	P27406 feline cali
9	267	9.2	668	1 COAT_FCVF6	P27404 feline cali
10	153.5	5.3	2193	1 POLG_HE71B	Q66478 h genome po
11	151.5	5.2	2193	1 POLG_HE71M	Q66479 h genome po
12	145.5	5.0	2194	1 POLG_HE701	P32537 h genome po
13	145	5.0	2193	1 POLG_CX16T	Q94f31 c genome po
14	144	5.0	855	1 POLG_HRV3	Q82081 human rhino
15	143	4.9	2206	1 POLG_POL3L	P03302 poliovirus
16	140.5	4.9	2206	1 POLG_POL32	P06209 poliovirus
17	131.5	4.5	2205	1 POLG_POL2W	P23069 p genome po
18	128.5	4.4	2193	1 POLG_CX16C	Q65900 c genome po
19	128.5	4.4	2207	1 POLG_POL2L	P06210 p genome po
20	124.5	4.3	2196	1 POLG_C05N	Q9ylj1 e genome po
21	123	4.2	2175	1 POLG_BOVEV	P12915 b genome po
22	122.5	4.2	862	1 POLG_EC16H	Q66790 echovirus 1
23	122.5	4.2	2209	1 POLG_POL1S	P03301 p genome po
24	121.5	4.2	2208	1 POLH_POL1M	P03300 p genome po
25	119	4.1	2164	1 POLG_HRV89	P07210 h genome po
26	119	4.1	2185	1 POLG_SVDVH	P16604 s genome po
27	119	4.1	2185	1 POLG_SVDVU	P13900 s genome po
28	119	4.1	2194	1 POLG_EC30B	Q9wn78 e genome po
29	118	4.1	832	1 POLG_HRV1A	P23008 human rhino
30	117.5	4.1	2206	1 POLG_POL1M	P03299 p genome po
31	116.5	4.0	2150	1 POLG_HRV2	P04936 h genome po
32	116	4.0	1200	1 HYAL_STRP	Q76536 strongyloce
33	116	4.0	2185	1 POLG_CXB5P	Q03053 c genome po

34	115.5	4.0	1229	1 P121_HUMAN	Q9y2n3 homo sapien
35	115	4.0	888	1 POL_SMRVH	P03364 squirrel mo
36	113.5	3.9	2184	1 POLG_EC01F	O91734 e genome po
37	113.5	3.9	2333	1 POLG_FMDV1	P03306 f genome po
38	113	3.9	2201	1 POLG_CXA9	P21404 c genome po
39	112.5	3.9	708	1 VM2_RS0VL	P11077 reovirus lt
40	112.5	3.9	2214	1 POLG_CXA24	P36290 c genome po
41	112	3.9	2179	1 POLG_HRV14	P03303 h genome po
42	111.5	3.9	2157	1 POLG_HRV1B	P12916 h genome po
43	111.5	3.9	2206	1 POLG_CXA21	P22055 c genome po
44	111	3.8	896	1 CYRB_MOUSE	P26955 mus musculu
45	110.5	3.8	2153	1 POLG_HRV16	Q82122 h genome po

ALIGNMENTS

RESULT 1

ID	COAT_SOUV3	STANDARD;	PRT;	546 AA.
AC	Q04542;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	01-OCT-1994 (Rel. 30, Last annotation update)			
DE	Coat protein (Capsid protein).			
OS	Southampton virus (serotype 3).			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;			
OC	Norovirus.			
OX	NCBI_TaxID=37129;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93142023; PubMed=8380940;			
RA	Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;			
RT	"Sequence and genome organization of a human small round-structured			
RT	(Norwalk-like) virus."			
RL	Science 259:516-519(1993).			
CC	-!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.			

CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@sib-sib.ch).
CC	-----
DR	EMBL; L07418; AAA92984.1; -
DR	PIR; B37491; B37491.
DR	InterPro; IPR004005; Calici_coat.
DR	InterPro; IPR008975; Viral_cap_coat.
DR	Pfam; PF00915; Calici_coat; 1.
KW	Coat protein; Glycoprotein.
FT	CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE 546 AA; 58774 MW; C82B2A85AD4B05EA CRC64;

Query Match	99.7%; Score 2887; DB 1; Length 546;
Best Local Similarity	99.6%; Pred.No. 4.8e-20;
Matches 544; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	1 MMASKADAPQADGASGAGQLVPEVNTADPLPMEPVAGTAVATAGVNMIDPVMNF 60
Db	1 MMASKADAPQADGASGAGQLVPEVNTADPLPMEPVAGTAVATAGVNMIDPVMNF 60
QY	61 VQSPQGETTISPNNTPGDILFDLQGLPHNPFSLHLSQMYNGWGNMVRVILLAGNAFSA 120
Db	61 VQSPQGETTISPNNTPGDILFDLQGLPHNPFSLHLSQMYNGWGNMVRVILLAGNAFSA 120
QY	121 GKIIIVCCVPPPGFTSSSLTIAQTLPFHVIADVRLTEPIEMPLEDRVNLVYHTNDNQPTMR 180
Db	121 GKIIIVCCVPPPGFTSSSLTIAQTLPFHVIADVRLTEPIEMPLEDRVNLVYHTNDNQPTMR 180

Matches 250; Conservative 76; Mismatches 190; Indels 48; Gaps 17;

QY 1 MMASKADQASADGAGAGQALVPEVNTADPLPMEPVAGFTTAVATAGOVNMDIPWVNF 60
 DB 1 MKMASNDAMPS--DGS--AANLVPEVNN--EVMALFVPGAAIAAPVAGQONVIDPWIRNF 56
 QY 61 VQSPQCEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGNMVRILLAGNAPSA 120
 DB 57 VQAPGCEFTVSPNAPGEILMSAPGLDLPVLSHLSRMVNGYAGGFEVQVILAGNAFTA 116
 QY 121 GKIIVCCVPPGFTSSLTIAQATLPHVIAVDTLEPIEMPLEDVRNLYHTND--NQPTM 179
 DB 117 GKVIFAAPVPEFTTEGLSPESQVTFMFIIVDVQRQLEPVLIPLPDVNRNFYHYNQANDSTL 176
 QY 180 RLVMCLYTPLRTGGSGNSDSFVAGRVLTAPSSDSFSLFVPPPTIEQKTRAFVNIPL 239
 DB 177 KLIALMTPLR--ANNAGDDVFTVSCVLTFRSPDFDFLVPPTVESRTKPTVPVLTV 234
 QY 240 QTLNSRFPFSLIQMILSPDASQVQFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQGA 299
 DB 235 EEMSRSRFPPILEKLYTGPSSAFVQVQNGRCRTDGLVLLGTTLQLSAVNFCNFRGDVTHIA 294
 QY 300 ----RTNLTEVDCKPMAFDS--PAPVGPDPF--GKCDHMRISKTPNNTSSGD---PMRS 350
 DB 295 GSHDYTNVTLASQNSWNSYDPTTEEIPAPLGTFDFVKG-----IQGLLTQITRADGSTRAHK 348
 QY 351 VSVQTNVQGVFPHLSIQDFDEVFNHPTGDY-----IGTIEWISQSTPPTGTDINL 400
 DB 349 ATVSTGSHVHTPKLGSVQFTDINN---DFQAGQNTKFTPVQVIGDHHQNEP---QQ 401
 QY 401 WEIPDYGSSLSQAANLAPVFPFGFGEALYVVSAPFGPNRSAPN--DVPCLLPQEVITH 459
 DB 402 WSLPNYSGRTGHNHVLAPAVAPTFFPGQQLFFRSTWPGCS--GYPNMNLDCLLPQEWVLH 459
 QY 460 FVSECAPTMGDAALLHVDPDTRNRLGELFKLYPGGYLTICVPGVAGPQO--LPLNGVFL 517
 DB 460 FYQEAAPQASDVALLRFPNPTGRVLFECCLKHSGYITV---AHTGPYDLVLPNGYFR 515
 QY 518 FVSWVSRYQLKPVGTASTARSRL 541
 DB 516 FDSWVQVFTYLLAPMNGNGTGRREAL 539

RESULT 3
 ID POLN RHDV STANDARD; PRT; 2344 AA.
 AC P27410;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Non-structural polyprotein [Contains: RNA-directed RNA polymerase
 (EC 2.7.7.48); Thiol protease P3C (EC 3.4.22.-); Helicase (2C like
 protein); Coat protein].
 DE Rabbit hemorrhagic disease virus (RHDV).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Lagovirus.
 OC NCBI_TaxID=11976;
 RN SEQUENCE FROM N.A.
 RP MEDLINE=91361557; PubMed=1840711;
 RA Meyers G, Wirblich C, Thiel H.-J.;
 RT "Rabbit hemorrhagic disease virus -- molecular cloning and nucleotide
 sequencing of a calicivirus genome.";
 RL Virology 184:664-676(1991).
 CC -/- FUNCTION: THE P3C CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE
 CC FOR CLEAVAGE AT CERTAIN Q/G SITES IN THE POLYPROTEIN.
 CC -/- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA)(N).
 CC -/- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

QY 181 LVCMLYTPLTGSGSGNSDSFVAGRVLTAPSSDSFSLFVPPPTIEQKTRAFVNIPLQ 240
 DB 181 LVCMLYTPLTGSGSGNSDSFVAGRVLTAPSSDSFSLFVPPPTIEQKTRAFVNIPLQ 240
 QY 241 TLSNRRFPFSLIQMILSPDASQVQFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQGA 300
 DB 241 TLSNRRFPFSLIQMILSPDASQVQFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQGA 300
 QY 301 TLNLTEDVCKPMAFDSAPVGPDPFGKCDHMRISKTPNNTSGDPMSRVSQVQNGVF 360
 DB 301 TLNLTEDVCKPMAFDSAPVGPDPFGKCDHMRISKTPNNTSGDPMSRVSQVQNGVF 360
 QY 361 VPHLSIQDFDEVFNHPTGDYIGTIEWISQSTPPTGTDINLMEIPDYGSSLSQAANLAPV 420
 DB 361 VPHLSIQDFDEVFNHPTGDYIGTIEWISQSTPPTGTDINLMEIPDYGSSLSQAANLAPV 420
 QY 421 FPGFGEALYVVSAPFGPNRSAPNDVPCLLPQEVITHFVSEQAPTMDGDAALLHVDPD 480
 DB 421 FPGFGEALYVVSAPFGPNRSAPNDVPCLLPQEVITHFVSEQAPTMDGDAALLHVDPD 480
 QY 481 TNRNLGELFKLYPGGYLTICVPGVAGPQQLPLNGVFLFVSWVSRYQLKPVGTASTARSR 540
 DB 481 TNRNLGELFKLYPGGYLTICVPGVAGPQQLPLNGVFLFVSWVSRYQLKPVGTASTARSR 540
 QY 541 LGVRR 546
 DB 541 LGVRR 546

RESULT 2
 ID COAT LORDV STANDARD; PRT; 539 AA.
 AC P54635;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Coat protein (Capsid protein).
 DE Lordsdale virus (Human enteric calicivirus).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 OC NCBI_TaxID=82658;
 RN SEQUENCE FROM N.A.
 RP MEDLINE=96005060; PubMed=7561776;
 RA Dingle K.E., Lambden P.R., Caul E.O., Clarke I.N.;
 RT "Human enteric Caliciviridae: the complete genome sequence and
 RT expression of virus-like particles from a genetic group II small
 RT round structured virus";
 RL J. Gen. Virol. 76:2349-2355(1995).
 CC -/- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; X86557; CAA60255.1;
 DR InterPro; IPR004005; Calici coat.
 DR InterPro; IPR008975; Viral_Cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 539 AA; 58775 MW; 5E5C63E7F2C5FD21 CRC64;
 Query Match 38.5%; Score 1114; DB 1; Length 539;
 Best Local Similarity 44.3%; Pred. No. 5.le-73;

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M67473; AAA47285.1; --
 DR PIR; A41039; RWRWRH.
 DR PDB; 1KHV; 16-JAN-02.
 DR PDB; 1KHW; 16-JAN-02.
 DR MEROPS; C24.001; --
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR000317; Peptidase_C24.
 DR InterPro; IPR000605; RNA helicase.
 DR InterPro; IPR007095; RNA pol DS PS.
 DR InterPro; IPR001205; RNA pol p3D.
 DR InterPro; IPR007094; RNA pol psvir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 DR Pfam; PF03510; Peptidase_C24; 1.
 DR Pfam; PF00680; RNA dep RNA_pol; 1.
 DR Pfam; PF00910; RNA helicase; 1.
 DR PRINTS; PR00916; 2CENDOPTASE.
 DR PRINTS; PR00918; CALICIVIRUSNS.
 DR SMART; SM00382; AAA; 1.
 KW Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;
 KW Thiol protease; Helicase; ATP-binding; Coat protein; 3D-structure.
 FT CHAIN ?508 7707 HELICASE (P2C-LIKE).
 FT CHAIN 1109 1251 PROTEASE P3C.
 FT CHAIN ?1497 71625 RNA-DIRECTED RNA POLYMERASE.
 FT CHAIN ? 2344 COAT PROTEIN.
 FT NP_BIND 522 529 ATP (POTENTIAL).
 FT ACT_SITE 1212 1212 PROTEASE (POTENTIAL).
 FT ACT_SITE 1227 1227 PROTEASE (POTENTIAL).
 SQ SEQUENCE 2344 AA; 257066 MW; 1454C248F81E9212 CRC64;

Query Match 11.3%; Score 328; DB 1; Length 2344;

Best Local Similarity 25.4%; Pred. No. 2e-15;
 Matches 141; Conservative 64; Mismatches 238; Indels 112; Gaps 24;

QY 5 SKDAPQSGADGAGAGQLVPEVNTADPLMEPVAGTTAV-----ATAG-----Q 48
 DB 1770 AAAPQGEAAGTATTASVPGTTDQ---MDPGVAVTAVTAAENSASIASIAGIGPPQQ 1826
 QY 49 VNMIDPWIWNNVQSPQSGFTSPNNTPGDILFDLQGLHNPFLSHLSQMYNGVGNMR 108
 DB 1827 VQQETWRTNFYND---VFTWSVADAPGSIILYVQHSPPNNPFTAVLSQMTAGWAGMQ 1883
 QY 109 VRILLAGNAPSAGKIIWCCVPPGF-TSSSLTIAQATLFPHVIAVDTLEPIEMPLEDVNR 167
 DB 1884 FRFIVAGSGVFGRLVRAVIPPGEIGPGLEVRQ---FPHVVIDARSLEPVTITMPDLRP 1940
 QY 168 VLYH-TNDNQPTWRLVCMLYTLPRTGGGSGNSDSFVAGRVLTAPSSDFSFLVPP--- 223
 DB 1941 NNYHTPDGFLVPTLVLYNNLINLPFGSGTS---AIQVTVETRESEDFEFVIMRAPSK 1997
 QY 224 TTEQKTRAFVNIPLQTLN-SRPSLIQGMILSPDASQVQVQFONGRCLIDQLLGTTP 282
 DB 1998 TVDSISPAGLLTTPVLTGVGNDRNNGQVIGLPVPGFSTC---NRHWNLSGYGWS 2054
 QY 283 ATSGQLFRVRGKIN-QGARTLNLTEY-DGKPFMAFDSP-----APVGFDPFGKCDWHMRIS 336
 DB 2055 PRFGDIDHRGSGASYSGSNATVLPQWYANAGSAIDNPISQVAPDGFDP-----MS 2105
 QY 337 KTPNNTSSGDPMRSVSQVNVQGVFPHLGSIQDFEVNHTPDGYIGTIEWISQPTPPT 396
 DB 2106 FVFPNGPG-----IPAGWVGFGAI-----WNSNSGAPNVT 2136
 QY 397 DINLWEIPDYSGSLSQANLAPPVPPG-----FGEA-----LVYFVSAFPGP 439
 DB 2137 TVQAYEL---GPATGAPGNLTPTNTSGAQTVAKSIYAVVTGTAQNAGLFWNASGIIST 2193

QY 440 NNRAPNDVPCLLPQSYITHFVSEQPTMGDAALLHYVDPTNRNLGBFKLYPGGYLTCV 499
 DB 2194 FNASAITTP--QPDRIVTTTCTPAAPFVGKNTPIMFA--SVVRETGDVNTAGS----- 2244
 QY 500 PNGV--GAGPOOLPL 512
 DB 2245 ANGTQYTGSGQLPLV 2259

RESULT 4

COAT_SMSV4
 ID_COAT_SMSV4 STANDARD; PRT; 703 AA.
 AC P36285;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Coat protein (Capsid protein).
 OS San Miguel sea lion virus (serotype 4) (SMSV 4).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OX NCBI_TaxID=36407;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92410750; PubMed=1529644;
 RA Neill J.D.;
 RT "Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lion virus: identification of conserved and non-conserved amino acid sequences among calicivirus capsid proteins."; Virus Res. 24:211-222(1992).
 CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M87482; AAA16220.1; --
 DR PIR; C48562; C48562.
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 482 482 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 703 AA; 77721 MW; C5DAD823B261073 CRC64;

Query Match 10.3%; Score 299; DB 1; Length 703;

Best Local Similarity 27.2%; Pred. No. 5.2e-14;
 Matches 108; Conservative 60; Mismatches 173; Indels 56; Gaps 17;

QY 4 ASKAPQSGADGAGAGQLVPE-----VNTADPLMEPVAGTTAVATAGVNMIDPWVNN 59
 DB 151 AESDGP-----GSAEIVTEEGTVVQQPAP-APTALATLATATCK-SVEQEMWT-- 199
 QY 60 FVQSPQSGFTSPNNTPGDILFDLQGLHNPFLSHLSQMYNGVGNMRVILLAGNAPS 119
 DB 200 -FFSYHTSINNWSVESQKILYSQALNPISINPLDHIKLYSTWSSGGIDVFTVSGGVF 258
 QY 120 AGKIIWCCVPCGFTS-SSLTIAQATLFPHVIAVDTLEPIEMPLEDVNRVLYHTNDNQPT 178
 DB 259 GKKLAALLVPPGVEPIESVSLQ---YPHVLFDAKQTEPVTITPDIRTLFHSMDTDT 315
 QY 179 MRLVCMLYTLPRLTGGGSGNSDSFVAGRVLTAPSSDFSFLVPPPTISQKTRAFVNP 238
 DB 316 TKLV---INPENGVENTKTSITVE-----TPSADFTFALLKPPGSLIKHGSISDLIP 368
 QY 239 LQTL--SNSRPFSLIQGMILSPDASQVQVQFONGRCL-IDGQLLG-TTPATSGQLFVRGK 294

Db 369 RNSAHMGNRWSTISGVSQPRV-----FQSNRHFDFDSTTTGWSTPYYPVBIKIQGK 423
 QY 295 INQAGARTLTLTEVDGKPPMAFDSAPVGFDFGKCDWHMRISKTPNNTSSGDPNRSVSQ 354
 Db 424 VGSNNKWFHVIDT---KALVPGIPDGWPDFTTIPD---ETKATNGNFSYGESYAGST- 475
 QY 355 TNVQGFVPHLGSIQDFEVPNHTGDYI-GTIEWISQP 390
 Db 476 -----TIKPNENSTHFKGTICGTLTVEIP 501

RESULT 5
 COAT_SMSV1 STANDARD; PRT; 702 AA.
 AC F362B4;
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Coat protein (Capsid protein).
 OS San Miguel sea lion virus (serotype 1) (SMSV 1).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OC NCBI_TaxID=36406;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92410750; PubMed=1529644;
 RA Neill J.D.;
 RT "Nucleotide sequence of the capsid protein gene of two serotypes of
 San Miguel sea lion virus: identification of conserved and non-
 conserved amino acid sequences among calicivirus capsid proteins.";
 RL Virus Res. 24:211-222(1992).
 CC -I- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 CC or send an email to license@isb-sib.ch).
 CC EMBL; M87481; AAA16217.1; -.
 DR PIR; A48562; A48562.
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 493 493 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 702 AA; 77850 MW; E6E5A58523DEE3D7 CRC64;

Query Match 10.1%; Score 293.5; DB 1; Length 702;
 Best Local Similarity 28.3%; Pred. No. 1.3e-13;
 Matches 102; Conservative 45; Mismatches 178; Indels 35; Gaps 12;
 QY 12 ADGAGAGOLVPEVNTADPLNMEPVAGPTTAVATAGQVNMID-PHVVNMFVQSGEFTI 70
 Db 153 SDGPGGADIVTEEGTVQOQVPQAQSALTTLAAASTGKTVDCEWTT---PFSHTAVNW 209
 QY 71 SPNPTFGDILFDLQGLPHNLPFLSHLSOMYNGWGNMVRILLAGNAFSAKIIIVCCVPP 130
 Db 210 STTEAGKTLFRAUSPELNPYLRHLSSTYTSWGGIDVRFVSGGVGGKLAALIVPP 269
 QY 131 GFTS-SSLTIAQATLPHVIAVDTLEPIEMPLEDNVNLHYTNDNQPTMRLVCMYTP 189
 Db 270 GIEPVESPTMLQ---YPHVLFDARQEPVIFIPDIRKTLXMSMDTDTTLRLVMYNEL 326
 QY 190 RTGGGNGSDSFVACRVLTAPSSDFSLFVPTTEIQKTRAFVTPNIPLOT--LSNRF 247
 Db 327 INPYEQSEPKS-SCSITVTRPSSDFTSLKPPGSLKXGIPSDLIIPRNSHWMGNRW 385

QY 248 PSLIQMILSPDASQVQVQNGRCL-IDQLLGTTPATSGOLFVRVKGKINQARTLNTE 306
 Db 386 WSTIDGFFVQPRV-----FQSNRHFDFDSTTTGWSTPYYPVBIKIQGK 440
 QY 307 VDGKPMFDSAPVGFDFGKCDWHMRISKTPNNTSSGDPNRSVSQTNVOGF 360
 Db 441 TE-----KSLVP-GLPD-----GWPDITPTAMTASNGNDYTVVAYEYRITNNGTHFKGF 488

RESULT 6
 COAT_FCVP4 STANDARD; PRT; 668 AA.
 AC P27405;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Coat protein (Capsid protein).
 OS Feline calicivirus (strain Japanese F4) (FCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OC NCBI_TaxID=11980;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91306470; PubMed=1853578;
 RA Tohya Y., Taniguchi Y., Takahashi E., Utagawa E., Takeda N.,
 RA Miyamura K., Yamazaki S., Mikami T.;
 RT "Sequence analysis of the 3'-end of feline calicivirus genome.";
 RL Virology 183:810-814(1991).
 CC -I- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 CC or send an email to license@isb-sib.ch).
 CC EMBL; D90357; BAA14371.1; -.
 DR PIR; B40481; VCMWFC.
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 399 399 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 668 AA; 73589 MW; 85BBDDB85804E503 CRC64;

Query Match 9.8%; Score 284; DB 1; Length 668;
 Best Local Similarity 24.7%; Pred. No. 5.8e-13;
 Matches 145; Conservative 74; Mismatches 246; Indels 122; Gaps 27;
 QY 23 PEVNTA-DPLNMEPVAGPTTAVATAGQVNMIDPHVVNMFVQSGEFTI-----SPNNT 75
 Db 133 PEQTAGVGVIAPSAQMSAADMASGKSVDSW-----EAFFSHTSVNWSSTSET 183
 QY 76 PGDILFDLQGLPHNLPFLSHLSOMYNGWGNMVRILLAGNAFSAKIIIVCCVPPGFTS- 134
 Db 184 QGKILFKQSLGFLNLPYLRHLSSTYTSWGGIDVRFVSGGVGGKLAALIVPPGVDPV 243
 QY 135 SSUTIAQATLPHVIAVDTLEPIEMPLEDNVNLHYTNDNQPTMRLVCMYTP 193
 Db 244 QSTSMQLQ---YPHVLFDARQEPVIFIPDIRKTLXMSMDTDTTLRLVMYNDLINPYA 300
 QY 194 GSGNDSFVACRVLTAPSSDFSLFVPTTEIQKTRAFVTPNIPLOTLSN-----R 246
 Db 301 NDSNSGCVIT--VETKPGDPFKHLLKPP-----GSVLTHGSIPLPKSSSLWIGNR 353

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; M86379; AAA79327.1; --
 CC EMBL; Z11536; CAA77636.1; --
 CC PIR; B43382; VCMWFF.
 CC InterPro; IPR004005; Calici_coat.
 CC InterPro; IPR008975; Viral_Cap_coat.
 CC Pfam; PF00915; Calici_coat; 1.
 CC Coat protein; Glycoprotein.
 CC CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 618 618 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC SEQUENCE 671 AA; 73441 MW; 33BEE86D8370D5E5 CRC64;

Query Match 9.3%; Score 268.5; DB 1; Length 671;
 Best Local Similarity 26.8%; Pred. No. 7.7e-12;
 Matches 110; Conservative 47; Mismatches 149; Indels 105; Gaps 18;

QY 10 QSADGAGAGQLVPEVNT-ADPLEMEFVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEF 68
 Db 124 EADGSGITA-----PEQGTWVGVIAPESQAMSTAADMATCKSVDSW-----EAPF 170

QY 69 TI-----SPNTPGDILDLQGLPHLPFLSHLSQWNGVGNMVRILLAGNAPSACK 122
 Db 171 SFHTSVNMSTSETQKILFQSLGPLNLYLHAKLYVAVSGSVIEVRSISGSGVFGGK 230

QY 123 IIVCCVPGFTS-SSLTIAQATLPPIHVIADRTLEPIEMPLEDRVNLVHTNDNQPTMRL 181
 Db 231 LAAIIVPPGVDPVOSTMQLQ---YPHVLFDARQVEPVIFCLPDLRSLYHLMSDITDTISL 287

QY 182 VCMLYTLP-RTGGSGNSDSFVAVAGRVLTAPSSDFSLFVLP-----TIEQKT 229
 Db 288 VIMVYNDLINPYANDANSSGCIVT--VETKPGDPDFKFLKPPGSMITHGSIPLSLIPKT 345

QY 230 RAFTVPNIPLOTLSNRPSPSLIQGMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLF 289
 Db 346 SSLMIGN-----RWYSDITDFVIRP-----FVQANRHFDFNQ---ETAGWSTFRF 388

QY 290 RVGRKINGCARTLNLTEVDGKPF---NAPDSPAPVPGVPDFKCDWHRISKTPNNTSSGD 346
 Db 389 R-----PISVITTEQNGAKLGIGVATDVIVP-GIPD-----GW-----PDTTIPGE 428

QY 347 PMRSVSVQTVNQGVPFVHLGSIQFDEVFVNHPTGDIYIGTIEMISQSPSTPPGTD 397
 Db 429 LI-----PAGDYAITNGTGNDDITATGYD 452

RESULT 9
 ID COAT FCVC6 STANDARD; PRT; 668 AA.
 AC P27404;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Coat protein (Capsid protein)
 DE Feline calicivirus (strain CFI/68 FIV) (FCV).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OC NCBI_TaxID=11979;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RX MEDLINE=91374597; PubMed=1716692;
 RA Neill J.D., Reardon I.M., Heinrichson R.L.;
 RT "Nucleotide sequence and expression of the capsid protein gene of
 RT feline calicivirus."
 RL J. Virol. 65:5440-5447 (1991).

[2]
 RN SEQUENCE FROM N.A.

RP Neill J.D.;
 RA Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.

CC -!- PTM: The N-terminus is blocked.

CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; M32819; AAA2925.1; --
 CC EMBL; U13992; AAC13993.1; --
 CC PIR; A40507; VCMWFF.
 CC InterPro; IPR004005; Calici_coat.
 CC InterPro; IPR008975; Viral_Cap_coat.
 CC Pfam; PF00915; Calici_coat; 1.
 CC Coat protein; Glycoprotein.
 CC CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC SEQUENCE 668 AA; 73550 MW; 9E52312108D441 CRC64;

Query Match 9.2%; Score 267; DB 1; Length 668;
 Best Local Similarity 31.7%; Pred. No. 9.9e-12;
 Matches 76; Conservative 33; Mismatches 97; Indels 34; Gaps 8;

QY 34 EPVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEFTI-----SPNTPGDILDLQGLP 87
 Db 145 EPNQMSSTAADMATCKSVDSW-----EAFSEHTSVNMSTSETQKILFQSLG 195

QY 88 HLNPLSHLSQWNGVGNMVRILLAGNAPSACKIIVCCVPGFTS-SSLTIAQATLP 146
 Db 196 LLNPVLTALKLYVAVSGSVDFRSISGSGVFGGLAAIIVVPGIDPVQSTMLQ---YP 252

QY 147 HVIADRTLEPIEMPLEDRVNLVHTNDNQPTMRLVCMLYTLP-RTGGSGNSDSFVAVAG 205
 Db 253 HYLFDARQVEPVIFSLPDLRSLYHLMSDITDTISLIMVYNDLINPYANDANSSGCIVT- 311

QY 206 RVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIPLOTLSN-----RPPSLIQGMILSP 258
 Db 312 -VETKPGDPDFKFLKPP-----GSMITHGSIPLSLIPKSSSLMIGNRFSWSDITDFVIRP 365

RESULT 10
 ID POLG HE71B STANDARD; PRT; 2193 AA.
 AC Q66478;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2
 DE (P1B); Coat protein VP3 (P1C); Coat protein P2C; Core protein P3A; Genome-
 DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Protease 3C)
 DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C).
 DE (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.46)].
 OS Human enterovirus 71 (strain BrCr) (Ev 71).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Enterovirus.
 OC NCBI_TaxID=69153;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=BRCR;
 RX MEDLINE=96434998; PubMed=8837884;
 RA Brown B.A., Pallansch M.A.;
 RT "Complete nucleotide sequence of enterovirus 71 is distinct from
 RT

74	NTP	-----GDILFDLQL-----	GPHLNPFSLHLSOMYNGVGNMR	108		
380	NVTNATSLMERL	RPVSAQAGKELCAVFRADPGRDGPQWQSTMLGQLCGYYTQWGSGL	439			
109	VRILLAGNAFSA	GIKIVCCPPGFTSSLLTAQATLPHVIAVDTLEPLEMLEDVNRV	168			
440	VTFMGTGSPMATCK	MLIAVTPPGGLPK-DRATAMLGTHVTDWDFLOSSVTLVPIWISNT	498			
169	LY--HTND--	--NQPTWRLICMLY-----	TPLRTGGGNSDSFVYAGRVLTAPSSDPSFL	218		
499	HYRAHARDGV	FDYITGLVSIWQYVYVPI-----	GAPNT-AVILA---LAAQAKNFTMK	550		
219	FL-----	VPPTIEQK-----	TRAPTVPNIPLQTLNSRPS--LIQMI	255		
551	LCKDTSHLQ	TSIQGDRVADVIESIGDSVSRALT-QALEPAPTQONTQVSSHRLOTGEV	609			
256	LSPDASQV	QFQN-----	GRCLIDGQLLGTTPATSGQLFVRGKINOGARTLNLT	307		
610	PALQAA	BIGASSNTSDSMETRCVLNSH--STAETTLDSF	PSRAGLVG-----EI	658		
308	DGKEFMA	FDGPVDFDKCDWHMRI	SKTPNNTSSGDPMRSVSVQTNVQGVFPHLGS	367		
659	D-LE	EGTTNPN-----	GYANWDIDI-----	YM	694	
368	QFDEVFNH	-----	PRGDVIGTIEWISQSP	TPPGTDINLWEIPDYGSSLQAA	NLAPPVF--	421
695	RFDAEFT	FAVCTPTGEVVPQL--LQYMFVPPGA-----	PKPESRESLAWQTATNPSFVVK	747		
422	---	PGFGEALVYFVS	APPGNNRSAPNDVPCLLPOEYITHEFVSQAP	PTMGDAALLHYVD	478	
748	LTDPP	-----	AQVSVPFMSPASAYQWFF-DGYPTFGS	---HKQE	782	
479	PD-----	TNRNLGEFKLYPGGYLTCVPNGVAGAPQOLPLNGVFLFV-----	SWVSIFY	526		
783	KDLEYGAC	PNMGMCTFSV-----	RTVGSKSKYPL-VVRIYMRMGHVRAMIPRPM	831		
527	Q-----	-----	LKPVGVTASTARSRLG	542		
832	RNQNLYFKAN	PNYAGNSIKPTGSRNAITL	GL	863		
RESULT 12						
POLG_HE701	STANDARD; PRT; 2194 AA.					
AC	P32537;					
DT	01-OCT-1993 (Rel. 27, Created)					
DT	01-NOV-1995 (Rel. 32, Last annotation update)					
DT	10-OCT-2003 (Rel. 42, Last annotation update)					
DE	Genome polypeptide [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)].					
OS	Human enterovirus 70 (strain J670/71) (EV 70).					
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;					
OC	Enterovirus					
OX	NCBI_Taxid=31915;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=91037960; PubMed=2172447;					
RA	Ryan M.D., Jenkins O., Hughes P.J., Brown A., Knowles N.J., Booth D.,					
RA	Minor P.D., Almond J.W.;					
RT	"The complete nucleotide sequence of enterovirus type 70: relationships with other members of the picornaviridae.";					
RL	J. Gen. Virol. 71:2291-2299(1990).					
CC	- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.					
CC	- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the poliovirus polypeptide. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.					
CC	- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +					
CC	- [RNA] (N).					

CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D00820; BAA1891.1; -
 CC PIR; A36253; GNNY87.
 CC HSP; P03300; IPOV.
 CC MEROPS; C03.UBA; -
 CC MEROPS; C03.UBB; -
 CC InterPro; IPR003593; AAA_ATPase.
 CC InterPro; IPR004004; Calici_pol_hel.
 CC InterPro; IPR009003; Cys_ser_trypsin.
 CC InterPro; IPR000199; pept_3C_picorn.
 CC InterPro; IPR000081; peptidase_C3.
 CC InterPro; IPR003338; Pico_P1A.
 CC InterPro; IPR002527; Pico_P2B.
 CC InterPro; IPR001676; Rhv.
 CC InterPro; IPR000605; RNA_helicase.
 CC InterPro; IPR007095; RNA_pol_DS_PS.
 CC InterPro; IPR001205; RNA_pol_P3D.
 CC InterPro; IPR007094; RNA_pol_PSVir.
 CC InterPro; IPR008975; Viral_cap_coat.
 CC Pfam; PF00548; Cys-protease_3C; 1.
 CC Pfam; PF02226; Pico_P1A; 1.
 CC Pfam; PF00947; Pico_P2A; 1.
 CC Pfam; PF01552; Pico_P2B; 1.
 CC Pfam; PF00073; rhv; 3.
 CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
 CC Pfam; PF00910; RNA_helicase; 1.
 CC PRINTS; PR00918; CALICIVIRUS.
 CC ProDom; PD001125; Cys_protease_3C; 1.
 CC ProDom; PD001306; Pico_P2A; 1.
 CC ProDom; PD001274; Pico_P2B; 1.
 CC SMART; SM00387; AAA; 1.
 CC PolyProtein; Coat protein; Core protein; Thiol protease; Myristate;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
 KW Lipoprotein.
 FT CHAIN 2 69 COAT PROTEIN VP4.
 FT CHAIN 70 319 COAT PROTEIN VP2.
 FT CHAIN 320 561 COAT PROTEIN VP3.
 FT CHAIN 562 871 COAT PROTEIN VP1.
 FT CHAIN 872 1014 CORE PROTEIN P2A.
 FT CHAIN 1015 1113 CORE PROTEIN P2B.
 FT CHAIN 1114 1443 CORE PROTEIN P2C.
 FT CHAIN 1444 1532 CORE PROTEIN P3A.
 FT CHAIN 1533 1554 GENOME-LINKED PROTEIN VP3.
 FT CHAIN 1555 1737 PICORNAIN 3C.
 FT CHAIN 1738 2194 RNA-DIRECTED RNA POLYMERASE P3D.
 FT LIPID 2 2 N-myristoyl glycine (in host) (By
 FT similarity).
 FT ACT SITE 1701 1701 PROTEASE (POTENTIAL).
 FT ACT SITE 1715 1715 PROTEASE (POTENTIAL).
 SQ SEQUENCE 2194 AA; 244590 MW; 15DBAE96EE06673C CRC64;
 Query Match 5.0%; Score 145.5; DB 1; Length 2194;
 Best Local Similarity 27.1%; Pred. No. 0.028;
 Matches 59; Conservative 29; Mismatches 103; Indels 27; Gaps 9;
 QY 32 PMEVPAGPTTAVATAGQVN-----MIDPWVNNFVQSPQ-----EFTISPNNTGDIILF 81
 DB 342 PAFPDSPTPEWHFGQVHSMLEIVQIESMWEINNVNDASGVERLVRVQISAQSDMDQLLF 401
 QY 82 ----DLQL-GPHLNPFLSHLSQWYNGVGNMVRILLAGNAFSAKKIIVCCVPPGFTSSS 136

Db 402 NIPLDIQLEGPLRNTLLGNISRYTHWSGSLEMTMFCGSEFTTKLICYTPPG-GSSP 460
 QY 137 LTIAQATLFPFHVIADVRLTEPIEMPLEDVRNVLYHTND-----NPTMRLVCLMYLTPLR 190
 Db 461 TDRMQAMLAHVWDFGLQSSITIIIPWISGSHYRMENTDAKAINANVGYVTCFMQTNLV 520
 QY 191 TGGSGNSDSFVAVRVTAPSSDFSF-LFLVPPPIEQ 227
 Db 521 APVGA--ADQCIVGMV--AAKDFNLRLMRDSPDIGQ 554
 RESULT 13
 POLG_CX16T
 ID POLG CX16T STANDARD; PRT; 2193 AA.
 AC Q9QF31;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2
 DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein
 DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-
 DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)
 DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].
 OS Coxsackievirus A16 (strain Taiwan/5079/98).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Enterovirus.
 CC NCBI_TaxID=231417;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21427116; PubMed=11536241;
 RA Yan J.-J., Su I.-J., Chen P.-F., Liu C.-C., Yu C.-K., Wang J.-R.;
 RT "Complete genome analysis of enterovirus 71 isolated from an outbreak
 RT in Taiwan and rapid identification of enterovirus 71 and
 RT coxsackievirus A16 by RT-PCR";
 RL J. Med. Virol. 65:331-339(2001).
 CC -1- FUNCTION: It is thought that the P2C protein attaches to vesicular
 CC membranes and is associated with viral RNA synthesis.
 CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
 CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
 CC -1- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the
 CC poliovirus polyprotein. In other picornavirus reactions Glu may be
 CC substituted for Gln, and Ser or Thr for Gly.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA)(N).
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF177911; AAD55085.1; -
 CC HSP; P03300; IPOV.
 CC InterPro; IPR003593; AAA_ATPase.
 CC InterPro; IPR004004; Calici_pol_hel.
 CC InterPro; IPR009003; Cys_ser_trypsin.
 CC InterPro; IPR000199; pept_3C_picorn.
 CC InterPro; IPR000081; peptidase_C3.
 CC InterPro; IPR003138; Pico_P1A.
 CC InterPro; IPR002527; Pico_P2B.
 CC InterPro; IPR001676; Rhv.
 CC InterPro; IPR000605; RNA_helicase.
 CC InterPro; IPR007095; RNA_pol_DS_PS.
 CC InterPro; IPR001205; RNA_pol_P3D.
 CC InterPro; IPR007094; RNA_pol_PSVir.

```
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00548; Cys-protease_3C; 1.
DR Pfam; PF02226; Pico_P1A; 1.
DR Pfam; PF00947; Pico_P2A; 1.
DR Pfam; PF01552; Pico_P2B; 1.
DR Pfam; PF00073; rhv; 3.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICVIRUSNS.
DR ProDom; PD001125; Cys_protease_3C; 1.
DR ProDom; PD001306; Pico_P2A; 1.
DR ProDom; PD001274; Pico_P2B; 1.
DR SMART; SM00382; AAA; 1.
KW Polypeptide; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
KW Lipoprotein.
FT CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 323 COAT PROTEIN VP2.
FT CHAIN 324 565 COAT PROTEIN VP3.
FT CHAIN 566 862 COAT PROTEIN VP1.
FT CHAIN 863 1012 CORE PROTEIN P2A.
FT CHAIN 1013 1111 CORE PROTEIN P2B.
FT CHAIN 1112 1440 CORE PROTEIN P2C.
FT CHAIN 1441 1526 CORE PROTEIN P3A.
FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1549 1731 PICORNAVIRIN 3C.
FT CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE.
FT LIPID 2 2 N-myristoyl glycine (in host) (By
similarity).
FT ACT_SITE 1695 1695 PROTEASE (POTENTIAL).
FT ACT_SITE 1709 1709 PROTEASE (POTENTIAL).
FT SEQUENCE 2193 AA; 243182 MW; 927839DB58F61E7F CRC64;

Query Match 5.0%; Score 145; DB 1; Length 2193;
Best Local Similarity 20.28; Pred. No. 0.031;
Matches 105; Conservative 171; Indels 184; Gaps 22;

QY 20 QLVPEVN---TADPLMEPVAGPTAVATAGQVNMIDPWVNNFVQSPQGEFTISNNWP 76
Db 373 ETILEVNNLKTNETPMQRLCFPV-----VQSKTGELCAAFRADP 413
QY 77 GDILFDLQGLHNPFLSHLSQYNGVGNMVRILLAGNAFAGKIIVCCVPPGFT--- 133
Db 414 G-----RDGPWQSTILQQLCRYTQMSGSLSEVTFMFAGSMATGKMLIATPPGGNVPA 467
QY 134 -----SSSLTI-----AQATLFPHVIVADVETL-----E 156
Db 468 DRITAMLGTHVWDFGLQSSVTLVVPVMSINTHYRAHARAGVFDYTTTGIITWYQTNVYV 527
QY 157 PIEMP-----LEDVR-----NVLVHTNDNQPTMLRVC 183
Db 528 PIGAPTAYIVALAAQDNFTMKLCKDTEIQGTANTQGDPIADIMIDQTVNNQVNRSLTA 587
QY 184 MLYTP-----LRTGGGNSDSFVAVGRVL-----TAPSS 213
Db 588 LQVLPTAADTEASSHRLGTGVVPALQAAETGASSNASDKNLIETRCVNLHHSTQETAIGN 647
QY 214 DFSFLFLV-----PPTIEQKTAFTVNPIDPOTLSNRP-SLIQGMILSPDASQVQVQF 267
Db 648 FFSRAGLVSIITWPTTGTQNTDGYVNWIDLMGYAQLRKCELFYMRFDAEFTFVVAKP 707
QY 268 NGRCLIDGOLL-----GTTTATSGQLFRVGRKKNQARTLNLTVDGKPFMAFDSFAP 320
Db 708 NGELV--PQLLQVMVVPQAPKPTSRDSPAQTATNPVS-FVKMTDPPAQVSVFPMSPAS 764
QY 321 V-----GPFDFGKCDWHRIS-----KTPNNSTSGDPMRSVSVQTVNQGVFPHLGSIQF 369
Db 765 AQWQFYDGYPTFGE---HLQANDLDYGQCPNNMGTFSIRTVTGTEKS-----PHSITLRV 816
QY 370 DEVFNHPTGDYGTIEWISQP-----STPP---GTDI 398
Db 817 YMRKIHVRA-----WIPRLRNQPYLFKTNPNYKNDI 849
```

```
RESULT 14
POLG HRV3 STANDARD; PRT; 855 AA.
AC Q82081;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2
DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D)] (Fragment).
OS Human rhinovirus 3 (HRV-3).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Rhinovirus.
OX NCBI_TaxID=44130;
RN [1] SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RP MEDLINE=97094977; PubMed=8939746;
RA Zhao R., Pevear D.C., Kremer M.J., Giranda V.L., Kofron J.A.,
RA Kuhn R.J., Rosemann M.G.;
RT "Human rhinovirus 3 at 3.0-A resolution.";
RL Structure 4:1205-1220(1996).
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U60874; AAB05616.1; -.
CC PDB; 1RHI; 12-MAR-97.
DR InterPro; IPR003138; Pico_P1A.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF02226; Pico_P1A; 1.
DR Pfam; PF00073; rhv; 3.
KW Polypeptide; Coat protein; Myristate; 3D-structure; Lipoprotein.
FT CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 331 COAT PROTEIN VP2.
FT CHAIN 332 567 COAT PROTEIN VP3.
FT CHAIN 568 855 COAT PROTEIN VP1.
FT LIPID 2 2 N-myristoyl glycine (in host).
FT HELIX 36 38
FT TURN 50 50
FT TURN 51 54
FT HELIX 57 57
FT STRAND 63 64
FT TURN 83 87
FT STRAND 88 89
FT TURN 90 94
FT STRAND 101 102
FT STRAND 103 105
FT TURN 113 115
FT STRAND 123 123
FT HELIX 126 128
FT TURN 129 129
FT STRAND 133 134
FT STRAND 138 140
FT TURN 142 143
FT STRAND 147 151
FT TURN 152 152
FT HELIX 153 155
FT TURN 156 157
FT HELIX 159 165
FT TURN 166 167
FT TURN 168 180
FT TURN 185 186
FT STRAND 188 197
FT TURN 198 198
FT STRAND 203 203
FT TURN 207 207
```

FT HELIX 213 216
FT HELIX 219 221
FT STRAND 223 224
FT TURN 225 226
FT TURN 231 232
FT STRAND 234 234
FT TURN 238 244
FT STRAND 245 245
FT HELIX 247 252
FT STRAND 255 259
FT TURN 260 262
FT STRAND 265 270
FT STRAND 279 279
FT STRAND 284 284
FT STRAND 287 298
FT TURN 301 302
FT STRAND 307 323
FT TURN 339 342
FT TURN 346 347
FT STRAND 354 354
FT TURN 357 358
FT STRAND 370 371
FT STRAND 373 373
FT TURN 374 379
FT STRAND 382 383
FT HELIX 395 398
FT STRAND 399 402
FT STRAND 404 404
FT TURN 405 405
FT STRAND 410 415
FT TURN 418 419
FT HELIX 421 423
FT TURN 424 425
FT HELIX 427 432
FT TURN 433 434
FT STRAND 435 439
FT STRAND 442 448
FT TURN 452 453
FT STRAND 455 455
FT STRAND 457 463
FT TURN 465 466
FT HELIX 473 476
FT TURN 477 478
FT STRAND 480 485
FT STRAND 491 496
FT STRAND 505 506
FT HELIX 511 513
FT STRAND 517 522
FT STRAND 526 527
FT TURN 530 531
FT STRAND 536 544
FT TURN 546 547
FT STRAND 549 553
FT STRAND 585 587
FT STRAND 590 590
FT TURN 599 600
FT STRAND 604 606
FT HELIX 614 616
FT TURN 617 617
FT STRAND 620 620
FT STRAND 623 624
FT HELIX 630 632
FT STRAND 633 633
FT HELIX 634 637
FT TURN 638 638
FT STRAND 642 651
FT TURN 655 656
FT HELIX 660 663
FT TURN 664 664
FT STRAND 666 669
FT STRAND 677 683
FT TURN 684 685

FT STRAND 686 702
FT STRAND 714 720
FT TURN 722 723
FT TURN 730 731
FT HELIX 733 736
FT STRAND 742 746
FT TURN 747 748
FT STRAND 750 755
FT STRAND 764 765
FT STRAND 770 771
FT STRAND 780 780
FT HELIX 784 786
FT STRAND 790 795
FT STRAND 804 822
FT STRAND 826 826
FT TURN 833 834
FT TURN 840 841
FT TURN 852 853
FT NON_TER 855 855
SQ SEQUENCE 855 AA; 94300 MW; E3F9C92CA2DA8AB1 CRC64;

Query Match 5.0%; Score 144; DB 1; Length 855;
Best Local Similarity 20.0%; Pred. No. 0.01;
Matches 101; Conservative 65; Mismatches 187; Indels 152; Gaps 20;

QY 2 MMASKDAPOSADGASGAGOLVPEVNTADPLME-----PVAGPTTAVATAGQVNM 51
Db LMVVEIAPLNA--PTGSSPTLPVTVTIAPMCTEFTGIRSRISIVPOGLPTTTLPGSQFLT 346
QY 52 ID-----PWVNVFVQSQEFTISPNNT----- 75
Db TDDRQSPSALPSYEPTPRIHIPGKVRNLEIIQVGTLPNMNNTGNDNVNLYLPLHADR 406
QY 76 PGDILFDLQ-----GPHLNPFLSHLSQMYNGVGNMVRILLIAGNAFSAKIIIVCCVPPG 131
Db 407 QNEIQFTKLYIGDGVFKTLLGEIAQYTHWSGSLRISLMTYTGALSSAKIILAYTPFG 466
QY 132 FTSSSLTIAQATLFPHVIAVRTLEPIEMPLEDVNRNLYHTNDNOFTWRLVCMLYTPLRT 191
Db 467 -TRGPEDRKKAMLGTHVVMDIGLQSTIVMTIPWTSQVQFR-----YTDPT 511
QY 192 GCGSGNSDSFVVAGRVLTAPSSD---PSFLELVPTTEQKTRFTVNPILQTLNSRF 247
Db 512 YTSAGYLSWYLTSLILPPQTSQVYLLSFISACP---DFKLRLMK-----DTQTSQT-- 562
QY 248 PSLIQMILSPDASQVQVQFQNGRCLID-----GOLLGTTPTAT--- 284
Db 563 DALTEG--LSDELEEVIVEKTKQTLASVSSGPKHTQSVPALETANETGATLTPRSDNVET 620
QY 285 -----SQQLFRVRGKINQAGARTLNLTVEVCGKPFMAFDPSPAPVGFDPFGKCDMMHRIK 337
Db 621 RTTYMHFNGSETDVSFLGRAA-CVHVTIKKNAAAGLDNHRKEGLFN---DWKINLSS 675
QY 338 TPNNTSSGDPMRSVSQVQNVQGFVPHLGSIQFDEVFNHPTGDIYGTIEWISQSPSPPGTD 397
Db 676 L-----VQLRKLELFT---YVRPDSBYT-----ILATASQP----- 704
QY 398 INLWEIPDYGSSLSQAANLAPVEFP 422
Db 705 ----EASSYSSNLTQAMVYVPPGAP 725

RESULT 15
POLG POL3L
ID POLG POL3L STANDARD; PRT: 2206 AA
AC P03302; Q84783; Q84784; Q84785; Q84786; Q84787; Q84788; Q84789;
AC Q84790; Q98592; Q98593; Q98594;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C, P3A; Genome-linked protein VPg; Picornain 3C
DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D

DE (EC 2.7.7.48)]].

OS Poliovirus type 3 (strains P3/Leon/37 and P3/Leon 12A[1]B).

OC Viruses; sRNA positive-strand viruses, no DNA stage; Picornaviridae;

NCBI_TaxID=12088;

OK

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=P3/Leon/37; PubMed=6324200;

RX MEDLINE=84170338; PubMed=6324200;

RA Stanway G., Hughes P.J., Mountford R.C., Reeve P., Minor P.D.,

RA Schild G.C., Almond J.W.;

RT "Comparison of the complete nucleotide sequences of the genomes of

RT the neurovirulent poliovirus P3/Leon/37 and its attenuated Sabin

RT vaccine derivative P3/Leon 12A[1]B.;"

RL Proc. Natl. Acad. Sci. U.S.A. 81:1539-1543(1984).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=P3/Leon 12A[1]B;

RX MEDLINE=8329239; PubMed=6310508;

RA Stanway G., Cann A.J., Hauptmann R., Hughes P.J., Clarke L.D.,

RA Mountford R.C., Minor P.D., Schild G.C., Almond J.W.;

RT "The nucleotide sequence of poliovirus type 3 Leon 12 alb: comparison

RT with poliovirus type 1.;"

RL Nucleic Acids Res. 11:5629-5643(1983).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 1-878.

RX MEDLINE=95120467; PubMed=7820548;

RA Grant R.A., Hiramath C.N., Filman D.J., Syed R., Andries K.,

RA Hogle J.M.;

RT "Structures of poliovirus complexes with anti-viral drugs:

RT implications for viral stability and drug design.;"

RL Curr. Biol. 4:784-797(1994).

RN [4]

RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1-878.

RA Hiramath C.N., Grant R.A., Filman D.J., Hogle J.M.;

RT "Binding of the antiviral drug wint51711 to the Sabin strain of type-3

RT poliovirus - structural comparison with drug-binding in rhinovirus-

RT 14.;"

RL Acta Crystallogr. D 51:473-489(1995).

CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN

CC O/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.

CC -1- CATALYTIC ACTIVITY: Selective cleavage of Gln-[Gly bond in the

CC poliovirus polypeptide. In other picornavirus reactions Glu may be

CC substituted for Gln, and Ser or Thr for Gly.

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +

CC {RNA} (N).

CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,

CC each of which is composed of one copy each of proteins VP1, VP2,

CC VP3, and VP4.

CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.

CC -1- MISCELLANEOUS: THE SEQUENCE OF STRAIN SABIN VACCINE P3/LEON/37 IS

CC SHOWN.

CC -1- MISCELLANEOUS: THE STRAIN SABIN VACCINE P3/LEON/37 IS THE

CC PROGENITOR OF THE STRAIN SABIN VACCINE P3/LEON 12A[1]B.

CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>

CC or send an email to license@ebi.ac.uk).

CC

CC EMBL; X01392; AAA46914.1; -.

CC EMBL; X00925; CAA25444.1; -.

CC PDB; 1PIV; 03-JUN-95.

CC PDB; 1PVC; 15-SEP-95.

CC PDB; 1VBA; 11-JUL-96.

CC PDB; 1VBB; 11-JUL-96.

CC PDB; 1VBC; 11-JUL-96.

CC PDB; 1VBE; 11-JUL-96.

CC MEROPS; C03.001; -.

DR MEROPS; C03.020; -.

DR InterPro; IPR004004; Calici_pol_hel.

DR InterPro; IPR009003; Cys_ser_trypsin.

DR InterPro; IPR000199; Pept_3C_picorn.

DR InterPro; IPR000081; Peptidase_C3.

DR InterPro; IPR003138; Pico_P1A.

DR InterPro; IPR002537; Pico_P2B.

DR InterPro; IPR001676; Rhv.

DR InterPro; IPR000605; RNA_helicase.

DR InterPro; IPR007095; RNA_pol_DS_PS.

DR InterPro; IPR001205; RNA_pol_P3D.

DR InterPro; IPR007094; RNA_pol_PSVir.

DR InterPro; IPR008975; Viral_cap_coat.

DR Pfam; PF00548; Cys-protease_3C; 1.

DR Pfam; PF02226; Pico_P1A; 1.

DR Pfam; PF00947; Pico_P2A; 1.

DR Pfam; PF01552; Pico_P2B; 1.

DR Pfam; PF00073; rhv; 3.

DR Pfam; PF00680; RNA_dep_RNA_pol; 1.

DR Pfam; PF00910; RNA_helicase; 1.

DR PRINTS; PD00918; CALICIVIRUSNS.

DR ProDom; PD001125; Cys_protease_3C; 1.

DR ProDom; PD001306; Pico_P2A; 1.

DR ProDom; PD001274; Pico_P2B; 1.

DR Polyprotein; Coat protein; Core protein; Transferase;

DR RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;

KW 3D-structure; Lipoprotein.

FT CHAIN 2 69

FT CHAIN 70 340

FT CHAIN 341 578

FT CHAIN 579 878

FT CHAIN 879 1027

FT CHAIN 1028 1124

FT CHAIN 1125 1453

FT CHAIN 1454 1540

FT CHAIN 1541 1562

FT CHAIN 1563 1745

FT CHAIN 1746 2206

FT LIPID 2 2

FT ACT SITE 1709 1709

FT ACT SITE 1723 1723

FT VARIANT 431 431

FT VARIANT 864 864

FT VARIANT 908 908

FT STRAND 4 8

FT STRAND 25 29

FT HELIX 36 38

FT TURN 50 50

FT HELIX 51 54

FT STRAND 57 57

FT TURN 63 64

FT TURN 80 81

FT STRAND 83 87

FT TURN 88 89

FT STRAND 90 94

FT STRAND 101 102

FT HELIX 103 105

FT TURN 113 115

FT STRAND 123 123

FT HELIX 126 128

FT TURN 129 129

FT STRAND 133 134

FT STRAND 138 141

FT TURN 142 143

FT STRAND 147 151

FT TURN 152 152

FT HELIX 153 155

FT TURN 156 157

FT HELIX 159 167

FT STRAND 168 180

FT TURN 185 186

FT STRAND 187 197

FT TURN 198 198

COAT PROTEIN VP4.

COAT PROTEIN VP2.

COAT PROTEIN VP1.

CORE PROTEIN P2A.

CORE PROTEIN P2B.

CORE PROTEIN P2C.

CORE PROTEIN P3A.

GENOME-LINKED PROTEIN VP6.

PICORNAIN 3C.

RNA-DIRECTED RNA POLYMERASE P3D.

N-myristoyl glycine (in host).

PROTEASE (POTENTIAL).

PROTEASE (POTENTIAL).

S -> F (IN P3/LEON 12A[1]B).

K -> R (IN P3/LEON 12A[1]B).

T -> A (IN P3/LEON 12A[1]B).


```
FT STRAND 203 209
FT TURN 208 216
FT HELIX 213 221
FT STRAND 219 224
FT STRAND 223 226
FT STRAND 226 236
FT TURN 235 242
FT STRAND 242 245
FT STRAND 245 248
FT TURN 249 252
FT HELIX 255 260
FT STRAND 263 267
FT TURN 268 270
FT STRAND 273 278
FT STRAND 287 287
FT TURN 289 291
FT STRAND 292 292
FT STRAND 295 307
FT TURN 308 310
FT STRAND 310 310
FT STRAND 314 331
FT STRAND 337 337
FT TURN 348 351
FT TURN 355 356
FT STRAND 363 363
FT TURN 366 367
FT STRAND 379 380
FT STRAND 382 382
FT TURN 383 383
FT HELIX 384 387
FT TURN 388 388
FT STRAND 391 392
FT STRAND 397 397
FT TURN 399 403
FT HELIX 405 408
FT STRAND 410 413
FT TURN 414 415
FT TURN 418 419
FT STRAND 423 426
FT TURN 429 431
FT TURN 433 437
FT HELIX 439 444
FT TURN 445 446
FT STRAND 447 451

Query Match 4.9%; Score 143; DB 1; Length 2206;
Best Local Similarity 21.4%; Pred. No. 0.043;
Matches 123; Conservative 62; Mismatches 220; Indels 170; Gaps 27;

QY 17 GAGQLVPEVNTADPLPMEPVAGFTTAVATAGV-NM-----IDPWVNNFVQSGQ---- 66
Db 349 GSNQVLTSDNHQSPCAI-PEPDVTPPIDIPGEVKNMELAEIDTWIPLNLESTKNTWDM 407
QY 67 -EFTISPNNTPGDILFDLQGLHNPFLSH-----LSQMYNGWGVNMRVRILLAGNAFSA 120
Db 408 YRVTLSDSADLSQFILCLSLSPASDPRLSHTMLGEVLNYYTHWAGSLKFTFLFCGMMAT 467
QY 121 GKIIVCCVPPGTSSSLTIAQATLPHVTADVTLEPIEMPLEDVNVLVH--TNDNQPT 178
Db 468 GKILVAYAPPG-AQPPTSKEAMLGTHVIWDLGLOSSCTMVVPWISNVTYRQTQDSFTE 526
QY 179 MRLVCMVLY-----TPLRTGGSGNSDSFVVAGRVLTAPSSDFSLFLVPPTIEQKRAFT 233
Db 527 GGYISMFYQTRIVVPLST-----PKSMSMLGFV--SACNDFSRLLRDFT----- 569
QY 234 VPNIPLQTLNSGRFPFSLIQGMILSPDASQVQFQNGRCLIDGQLLGTTPATSGQLFRVRG 293
Db 570 -----HISQALPQIEDLI-----S 585
QY 294 KINQAGARTLNL-TEVDGKPFMAFDSPAPVFPDFGKCDWHMR--ISKTPNNTSSGDPMR- 349
Db 586 EVAQAGALTLSLPKQDQSLPDTKASGPA-----HSKEVPALTAVETGATNPLAP 633
```

```
QY 350 SVSVQT-----NVQGFVPHLSIQDFEVNHPHTGDYICTIEMI SQSTPPGTD 397
Db 634 SDTVQTRHVQRRSRSESTIESFFARGACVAIIIEVDN-----EQPTTAAQKL 680
QY 398 INLWEIPDYGSLSQAANLAPPVPPGGEALVYFVSA-FPGPNRSAPNDVPCLLPQEY 456
Db 681 FAMMRI-TYKDTVQLRRKLEFFTY-SRFDMEFTFVVVTANFNANNGHALNQV-----Y 731
QY 457 ITHFVSEQAPTMGDAALLHYVDP-----DTNRNLGEFKLYPGGYLTCVPNGVGAGPQ 508
Db 732 QIMYIPPGCAPT-----PKSWDDYTWTQTSNPSIFITY-----GAAPA 768
QY 509 QLPLNGVFLFVSWVSRYQ-LKPVGTASTARSRLG 542
Db 769 RISVPYVGLANAY-SHFYDGFAPKVLKTDANDQIG 802

Search completed: June 1, 2004, 13:47:12
Job time : 8.59597 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:35:27 ; Search time.31.7889 Seconds
(without alignments)
5419.275 Million cell updates/sec

Title: US-09-926-799-3
Perfect score: 2895
Sequence: 1 MMASKDAPQSDAGSAGAGQ.....QLKPVGTASTARSLGVRR 546

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL_25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2895	100.0	546	12 Q8JW44	Q8JW44 norwalk-lik
2	2868	99.1	544	12 Q8VA27	Q8VA27 human calic
3	2865	99.0	546	12 Q9IV47	Q9IV47 human calic
4	2231.5	77.1	544	12 Q9DU46	Q9DU46 chiba virus
5	2228.5	77.0	544	12 Q9QTE7	Q9QTE7 chiba virus
6	2215.5	76.5	544	12 Q9I185	Q9I185 human calic
7	2210.5	76.4	544	12 Q9I8C5	Q9I8C5 norwalk-lik
8	2208	76.3	543	12 Q9IV43	Q9IV43 human calic
9	2203.5	76.1	544	12 Q9IEJ2	Q9IEJ2 human calic
10	2202	76.1	543	12 Q9I8B9	Q9I8B9 norwalk-lik
11	2201.5	76.0	544	12 Q9I8D1	Q9I8D1 norwalk-lik
12	2199.5	76.0	544	12 Q9IV36	Q9IV36 human calic
13	2179.5	75.3	544	12 Q9IV41	Q9IV41 human calic
14	2178	75.2	543	12 Q8UXJ0	Q8UXJ0 norwalk-lik
15	2159.5	74.6	531	12 Q8V7J5	Q8V7J5 norwalk-lik
16	2070	71.5	540	12 Q9YS14	Q9YS14 norwalk vir

17	2058	71.1	540	12 Q8BCA5	Q8BCA5 human calic
18	2030.5	70.1	530	12 Q838B4	Q838B4 norwalk vir
19	2029.5	70.1	530	12 Q9QT39	Q9QT39 norwalk-lik
20	2003.5	69.2	539	12 Q8JVV5	Q8JVV5 norwalk-lik
21	1993.5	68.9	530	12 Q83876	Q83876 norwalk vir
22	1991.5	68.8	517	12 Q9JH41	Q9JH41 norwalk-lik
23	1986.5	68.6	541	12 Q9IV42	Q9IV42 human calic
24	1985.5	68.6	545	12 Q8VA02	Q8VA02 human calic
25	1981	68.4	544	12 Q66418	Q66418 desert shie
26	1970.5	68.1	543	12 Q8BCA3	Q8BCA3 human calic
27	1964.5	67.9	545	12 Q9I8C8	Q9I8C8 norwalk-lik
28	1959.5	67.7	545	12 Q9WI82	Q9WI82 norwalk-lik
29	1953.5	67.5	545	12 Q9I8C2	Q9I8C2 norwalk-lik
30	1951.5	67.4	545	12 Q9IV45	Q9IV45 human calic
31	1947.5	67.3	539	12 Q9IV48	Q9IV48 human calic
32	1932.5	66.8	545	12 Q9IH12	Q9IH12 human calic
33	1298	44.8	522	12 Q8B558	Q8B558 bovine ente
34	1297	44.8	522	12 Q8B4Y9	Q8B4Y9 norwalk-lik
35	1290	44.6	522	12 Q8B4Z2	Q8B4Z2 norwalk-lik
36	1288	44.5	522	12 Q8B4Z5	Q8B4Z5 norwalk-lik
37	1286	44.4	520	12 Q7TBK7	Q7TBK7 bovine ente
38	1286	44.4	522	12 Q8V629	Q8V629 norwalk-lik
39	1285	44.4	522	12 Q8V628	Q8V628 norwalk-lik
40	1282	44.3	520	12 Q7TBL1	Q7TBL1 bovine ente
41	1225	42.3	556	12 Q9PY75	Q9PY75 human calic
42	1225	42.3	556	12 Q9I7V6	Q9I7V6 norwalk-lik
43	1222.5	42.2	519	12 Q9YQ22	Q9YQ22 bovine cali
44	1218	42.1	556	12 Q9I7V9	Q9I7V9 norwalk-lik
45	1185	40.9	535	12 Q9I5C6	Q9I5C6 human calic

ALIGNMENTS

RESULT 1

Q8JW44 Q8JW44 PRELIMINARY; PRT; 546 AA.
AC Q8JW44;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22; Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Capsid protein.
GN CAPSID.
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Funabashi258;
RA Katayama K., Takeda N., Natori K.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Funabashi258;
RA Natori K., Takeda N.;
RT "Genetic and antigenic relationship among Norwalk-like viruses.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB078335; BAC05516.1; ...
DR InterPro; IPR004605; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 546 AA; 58834 MW; 1965F054E2C481E6 CRC64;

Query Match 100.0%; Score 2895; DB 12; Length 546;
Best Local Similarity 100.0%; Pred. No. 3.2e-231;
Matches 546; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MMASKDAPQSDAGSAGAGQLVPEVNTADPLPMEPVAGTTAVATAGVNMIDPVMVNF 60
Db 1 MMASKDAPQSDAGSAGAGQLVPEVNTADPLPMEPVAGTTAVATAGVNMIDPVMVNF 60
QY 61 VQSQGEFTISPNTPGDILFDLQLPHLNPFLSHLSQWNGVGNMVRVILLAGNAFSA 120

[illegible]

	Query Match	99.0%	Score 2865;	DB 12;	Length 546;
	Best Local Similarity	98.9%;	Prod. No. 9.7e-229;		
	Matches 540;	Conservative	3; Mismatches	3; Indels	0; Gaps
QY	1	MMWASKDAPQADGASGAGQGVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPMTVNNF	60		
DB	1	MMWASKGAPQADGSSGAGQGVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPMTVNNF	60		
QY	61	VQSPQGEFTTISPNNTPGDILFDLQLGPHLNPFSLHSQMVNGHWGNMVRILLAGNAFSA	120		
DB	61	VQSPQGEFTISPNNTPGDILFDLQLGPHLNPFSLHSQMVNGHWGNMVRILLAGNAFSA	120		
QY	121	GKTIIVCCVPGFTSSSLTIAQATILFPHVIADVATLEPIEMPLEDVRNVLVYHTNDNQFTMR	180		

```

Db      121  GKIIVCCVPPGFTSSSLTIAQATLFPVIAVRLTLEPIEMPLEDRVRLVHTNDSTQPTMR 180
QY      181  LVCMLYTLRTGGSGNSDSFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIPLQ 240
Db      181  LVCMLYTLRTGGSGNSDSFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIPLQ 240
QY      241  TILSNRFPPLIQMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQGAR 300
Db      241  TILSNRFPPLIQMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQGAR 300
QY      301  TLNLTEVDGKPFMAFDPAPVGFDFGKCDWMNRISKTNNNTSSGDPMSVSQTNVQGF 360
Db      301  TLNLTEVDGKPFMAFDPAPVGFDFGKCDWMNRISKTNNNTSSGDPMSVSQTNVQGF 360
QY      361  VPHLSIQDFDEVNHTPGDYIGTIEWISQFSTPPGTDINLWEIPDYGSSLSQAANLAPV 420
Db      361  VPHLSIQDFDEVNHTPGDYIGTIEWISQFSTPPGTDINLWEIPDYGSSLSQAANLAPV 420
QY      421  FPPGFEALVYFVSAPFPNNRSAPNDVPCLLPQEVITHFVSEQAPTMGDALLHYVDP 480
Db      421  FPPGFEALVYFVSAPFPNNRSAPNDVPCLLPQEVITHFVSEQAPTMGDALLHYVDP 480
QY      481  TNRNLGEFKLYPGGYLTCVPGNGVAGPQQLPLNGVFLFVSWSRFYQLKPVGTASTARSR 540
Db      481  TNRNLGEFKLYPGGYLTCVPGNGVAGPQQLPLNGVFLFVSWSRFYQLKPVGTASTARSR 540
QY      541  LGVRR 546
Db      541  LGVRR 546

```

RESULT 4

```

Q9DU46
ID Q9DU46 PRELIMINARY; PRT; 544 AA.
AC Q9DU46;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Chiba virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=99565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Chiba 407/1987/JP;
RA Someya Y., Takeda N., Miyamura T.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Chiba 407/1987/JP;
RX MEDLINE=20569531; PubMed=11118371;
RA Someya Y., Takeda N., Miyamura T.;
RT "Complete nucleotide sequence of the Chiba virus genome and functional
expression of the 3C-like protease in Escherichia coli.";
RL Virology 278:490-500(2000).
DR EMBL; AB042808; BAB18267.1;
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58351 MW; 1DB7A4B81FEB246F CRC64;

```

```

Query Match 77.1%; Score 2231.5; DB 12; Length 544;
Best Local Similarity 74.7%; Pred. No. 3.1e-176;
Matches 407; Conservative 59; Mismatches 78; Indels 1; Gaps 1;
QY 1 MMASKDAPQSDAGSAGAGQLVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPWINNF 60
Db 1 MMASKDAPQSDAGSAGAGQLVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPWINNF 60
QY 61 VOSPOGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGVGNMVRVVLGNAPFA 120
Db 61 VOSPOGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGVGNMVRVVLGNAPFA 120

```

```

Db      61  VOAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGVGNMVRVVLGNAPFA 120
QY      121  GKIIVCCVPPGFTSSSLTIAQATLFPVIAVRLTLEPIEMPLEDRVRLVHTNDSTQPTMR 180
Db      121  GKIIVCCVPPGFTSSSLTIAQATLFPVIAVRLTLEPIEMPLEDRVRLVHTNDSTQPTMR 180
QY      181  LVCMLYTLRTGGSGNSDSFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIPLQ 240
Db      181  LVCMLYTLRTGGSGNSDSFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIPLQ 240
QY      241  TILSNRFPPLIQMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQGAR 300
Db      241  TILSNRFPPLIQMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQGAR 300
QY      301  TLNLTEVDGKPFMAFDPAPVGFDFGKCDWMNRISKTNNNTSSGDPMSVSQTNVQGF 360
Db      301  TLNLTEVDGKPFMAFDPAPVGFDFGKCDWMNRISKTNNNTSSGDPMSVSQTNVQGF 360
QY      361  VPHLSIQDFDEVNHTPGDYIGTIEWISQFSTPPGTDINLWEIPDYGSSLSQAANLAPV 420
Db      361  VPHLSIQDFDEVNHTPGDYIGTIEWISQFSTPPGTDINLWEIPDYGSSLSQAANLAPV 420
QY      421  FPPGFEALVYFVSAPFPNNRSAPNDVPCLLPQEVITHFVSEQAPTMGDALLHYVDP 480
Db      421  FPPGFEALVYFVSAPFPNNRSAPNDVPCLLPQEVITHFVSEQAPTMGDALLHYVDP 480
QY      481  TNRNLGEFKLYPGGYLTCVPGNGVAGPQQLPLNGVFLFVSWSRFYQLKPVGTASTARSR 540
Db      481  TNRNLGEFKLYPGGYLTCVPGNGVAGPQQLPLNGVFLFVSWSRFYQLKPVGTASTARSR 540
QY      541  LGVRR 545
Db      541  LGVRR 544

```

RESULT 5

```

Q9QTE7
ID Q9QTE7 PRELIMINARY; PRT; 544 AA.
AC Q9QTE7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Chiba virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=99565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Chiba 407;
RA Natori K., Suzuki K., Yamakawa Y., Tatsumi M., Sakae K., Kobayashi S.,
Shinozaki K., Ishiko H., Miyamura T., Takeda N.;
RT "Expression and self-assembly of capsid proteins of the Chiba virus, a
genetically distinct Norwalk-like virus.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022679; BAA82106.1;
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58323 MW; ABD1C1FC4F93D872 CRC64;

```

```

Query Match 77.0%; Score 2228.5; DB 12; Length 544;
Best Local Similarity 74.5%; Pred. No. 5.5e-176;
Matches 406; Conservative 59; Mismatches 79; Indels 1; Gaps 1;
QY 1 MMASKDAPQSDAGSAGAGQLVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPWINNF 60
Db 1 MMASKDAPQSDAGSAGAGQLVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPWINNF 60
QY 61 VOSPOGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGVGNMVRVVLGNAPFA 120
Db 61 VOSPOGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGVGNMVRVVLGNAPFA 120

```

Wed Jun 2 09:13:28 2004

us-09-926-799-3.rspt

121 GKIIICVPPGFTSSLTIAQATLPHVIADRTLEPIEMPLEDVRNVLXHTNDNQPTMR 180
 121 GKIIICVPPGFTSSLTIAQATLPHVIADRTLEPIEMPLEDVRNVLXHTNDNQPTMR 180
 181 LVCMLYTPURTGGSGNSDSFVAVGRVLTAPSSDFSLFVLPVPTTEOKTRAFVNIPLQ 240
 181 LVCMLYTPURTGGSGNSDSFVAVGRVLTAPSSDFSLFVLPVPTTEOKTRAFVNIPLK 240
 241 TILNSRFPSSLIOGMLISPDASQVQFQNGRCILIDGQLTTPATSGQLFRVRGKINQAR 300
 241 TILNSRFPSSLIOGMLISPDQTNQVQFQNGRCITIDGQLTTPVSVQSLCKFRGRTSQ 300
 241 YLSNSRIPNPIEGMSLSPDQTNQVQFQNGRCITIDGQLTTPVSVQSLCKFRGRTSQ 300
 301 TILNTEVDGKPFMAFDPSPAPVFPDFGKCDHMRISKTPNNTSSGDPMRSVSVQTNVQGF 360
 301 YLSNSRIPNPIEGMSLSPDQTNQVQFQNGRCITIDGQLTTPVSVQSLCKFRGRTSQ 300
 301 VLNLTELDGSPFMGFCAPAPAGFPDGLGSCDWHIEMSKIPNSSTQNNPIVTSVKNSQOF 360
 361 VPHLSITLDDNVS-SGGDYIGTIQWTSPPSDGGANTFWKIPDYGSSLAASQLAPAV 420
 361 VPHLSITLDDNVS-SGGDYIGTIQWTSPPSDGGANTFWKIPDYGSSLAASQLAPAV 419
 421 FPGFGEALYFVSAPFGPNRNSAPNDVPCLLPQEIYTHFVSEQAPTMGDALHLYVDPD 480
 420 YPGFNEVIYFVMASTPGPNQSGSNLVPCLLPQEIYTHFVSEQAPTMGDALHLYVDPD 479
 481 TNRNLGFEKLYPGGYLTCVNPNGVAGPQPLNGVFLFVSWYSRFFYQLKPVGTASTARS 540
 480 TNRNLGFEKLYPGGYLTCVNPSSSTGPGQLPLDGVFVAFASWSRFFYQLKPVGTAGPAR 539
 541 LGVRR 545
 540 LGVRR 544

RESULT 6

Q91185 Q91185 PRELIMINARY; PRT; 544 AA.
 AC Q91185;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Capsid protein.
 OS Human calicivirus HU/NLV/Koblentz/433/2000/DE.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 OX NCBI_TaxID=165252;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HU/NLV/Koblentz 433/2000/DE;
 RA Kuenkel U., Schreier E.;
 RT "Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF394960; AAK72048.1;
 DR InterPro; IPR004005; Calici.coat.
 DR InterPro; IPR008975; Viral.Cap.coat.
 DR Pfam; PF00915; Calici.coat.1;
 DR PIR; S58222 MW; 1B39A595B733A6A8 CRC64;
 SQ SEQUENCE 544 AA; 58222 MW; 1B39A595B733A6A8 CRC64;

Query Match 76.5%; Score 2215.5; DB 12; Length 544;
 Best Local Similarity 74.1%; Pred. No. 6.6e-175;
 Matches 404; Conservative 60; Mismatches 80; Indels 1; Gaps 1;

QY 1 MMASKADPQSDAGAGAGQLVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPWINNF 60
 DB 1 MMASKADPQSDAGAGAGQLVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPWINNF 60
 QY 61 VSPQCEFTISPNPTGDLFDLQGLPHLNPFLSHLSQMYNGVGNMVRVILLAGNAPSA 120
 DB 61 VQAPQCEFTISPNPTGDLFDLQGLPHLNPFLSHLSQMYNGVGNMVRVILLAGNAPSA 120
 QY 121 GKIIICVPPGFTSSLTIAQATLPHVIADRTLEPIEMPLEDVRNVLXHTNDNQPTMR 180
 DB 121 GKIIICVPPGFTSSLTIAQATLPHVIADRTLEPIEMPLEDVRNVLXHTNDNQPTMR 180

QY 181 LVCMLYTPURTGGSGNSDSFVAVGRVLTAPSSDFSLFVLPVPTTEOKTRAFVNIPLQ 240
 DB 181 LVCMLYTPURTGGSGNSDSFVAVGRVLTAPSSDFSLFVLPVPTTEOKTRAFVNIPLK 240
 QY 241 TILNSRFPSSLIOGMLISPDASQVQFQNGRCILIDGQLTTPATSGQLFRVRGKINQAR 300
 DB 241 YLSNSRIPNPIEGMSLSPDQTNQVQFQNGRCITIDGQLTTPVSVQSLCKFRGRTSQ 300
 QY 301 TILNTEVDGKPFMAFDPSPAPVFPDFGKCDHMRISKTPNNTSSGDPMRSVSVQTNVQGF 360
 DB 301 VLNLTELDGSPFMGFCAPAPAGFPDGLGSCDWHIEMSKIPNSSTQNNPIVTSVKNSQOF 360
 QY 361 VPHLSITLDDNVS-SGGDYIGTIQWTSPPSDGGANTFWKIPDYGSSLAASQLAPAV 420
 DB 361 VPHLSITLDDNVS-SGGDYIGTIQWTSPPSDGGANTFWKIPDYGSSLAASQLAPAV 419
 QY 421 FPGFGEALYFVSAPFGPNRNSAPNDVPCLLPQEIYTHFVSEQAPTMGDALHLYVDPD 480
 DB 420 YPGFNEVIYFVMASTPGPNQSGSNLVPCLLPQEIYTHFVSEQAPTMGDALHLYVDPD 479
 QY 481 TNRNLGFEKLYPGGYLTCVNPNGVAGPQPLNGVFLFVSWYSRFFYQLKPVGTASTARS 540
 DB 480 TNRNLGFEKLYPGGYLTCVNPSSSTGPGQLPLDGVFVAFASWSRFFYQLKPVGTAGPAR 539
 QY 541 LGVRR 545
 DB 540 LGVRR 544

RESULT 7

Q918C5 Q918C5 PRELIMINARY; PRT; 544 AA.
 AC Q918C5;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Capsid protein.
 GN ORF2.
 OS Norwalk-like virus NLV/Baltimore/277/1993/US.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 OX NCBI_TaxID=171836;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HU/NLV/Baltimore/277/1993/US;
 RX MEDLINE=97193806; PubMed=9041391;
 RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
 RT "A one-tube method of reverse transcription-PCR to efficiently amplify a 3-kilobase region from the RNA polymerase gene to the poly(A) tail of small round-structured viruses (Norwalk-like viruses).";
 RL J. Clin. Microbiol. 35:570-577(1997).
 RN [2]
 RP SEQUENCE OF 100-192 FROM N.A.
 RC STRAIN=HU/NLV/Baltimore/277/1993/US;
 RX MEDLINE=98071277; PubMed=9407386;
 RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K., Seto Y., Monroe S.S., Glass R.I.;
 RT "Correlation of patient immune responses with genetically characterized small round-structured viruses involved in outbreaks of nonbacterial acute gastroenteritis in the United States, 1990 to 1995.";
 RL J. Med. Virol. 53:372-383(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HU/NLV/Baltimore/277/1993/US;
 RX MEDLINE=20266071; PubMed=10804147;
 RA Ando T., Noel J.S., Fankhauser R.L.;
 RT "Genetic classification of 'Norwalk'-like viruses.";
 RL J. Infect. Dis. 181:S336-S348(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HU/NLV/Baltimore/277/1993/US;

RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414404; AAL12865.1; -;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58208 MW; D5170FA4B958A672 CRC64;

Query Match 76.4%; Score 2210.5; DB 12; Length 544;
Best Local Similarity 73.9%; Pred. No. 1.7e-174;
Matches 403; Conservative 61; Mismatches 80; Indels 1; Gaps 1;

QY 1 MMASKADPQSDAGSAGAGOLVPEVNTADPLMEPVAGPTTAVATAGVNMIDPVIWNNF 60
DB 1 MMASKADPQSDAGSAGAGOLVPEVNTADPLMEPVAGPTTAVATAGVNMIDPVIWNNF 60
QY 61 VQAPQGEFTISPNNTPGDIPLDQLGPHLNPFLSHLSQMYNGWGMVRVILLAGNAFSA 120
DB 61 VQAPQGEFTISPNNTPGDIPLDQLGPHLNPFLSHLSQMYNGWGMVRVILLAGNAFSA 120
QY 121 GKIIICVCPVPGTSSSLTIAQATLFPHVIAVRLTPEIEMPLEDVRNVLVYHTNDNQPTWR 180
DB 121 GKIIICVCPVPGTSSSLTIAQATLFPHVIAVRLTPEIEMPLEDVRNVLVYHTNDNQPTWR 180
QY 181 LVCMLYTPLRTGCGSGNSDSFVAGRVLTAPSDSFLVPPTEIQKTRAFVVPNIPLQ 240
DB 181 LVCMLYTPLRTGCGSGNSDSFVAGRVLTAPSDSFLVPPTEIQKTRAFVVPNIPLQ 240
QY 241 TLNSRPFSLIQMILSPDASQVQFQNGRCLIDGQLGTTTATSGQLFRVRGKINQGAR 300
DB 241 TLNSRPFSLIQMILSPDASQVQFQNGRCLIDGQLGTTTATSGQLFRVRGKINQGAR 300
QY 301 TLNLTVDGKPMAPFSPAPVGPDPFGKCDWHMRISKTPNNTSSGDPMSRSVQTNVQGF 360
DB 301 VLNLTELDCGSPFPAAPAGPDLGSCDWHIEMSKIPNSQNNPIVNSVKPNSQGF 360
QY 361 VPHLSGIQFDEVNHTPDYIGTIEWISQPS-TPGTDINLWEIPDYGSSLSQANLAPPV 420
DB 361 VPHLSGITLDENVIS-SCGDYIGTIEWISQPS-TPGTDINLWEIPDYGSSLSQANLAPPV 420
QY 421 FPPGFGALVYFVSAPFPNRRSAPNDVPCLLPQEIYTHFVSEQATMGDAALLHYVDP 479
DB 420 YPPGFNEVYFVWASIPGNSQSPNLVPCLLPQEIYTHFVSEQATMGDAALLHYVDP 479
QY 481 TNRNLGEFKLYPGGYLTCVPSNGVAGPQQLPLNGVFLFVSWSRFYQLKPVGTASTARS 540
DB 480 TNRNLGEFKLYPGGYLTCVPSNGVAGPQQLPLNGVFLFVSWSRFYQLKPVGTASTARS 540
QY 541 LGVRR 545
DB 540 LGVRR 544

RESULT 8
Q91V43 ID Q91V43 PRELIMINARY; PRT; 543 AA.
AC Q91V43
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Musgrove/89/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=122918;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=HU/NLV/Musgrove/89/UK;
RC MEDLINE=20404883; PubMed=10949950;
RX Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RL 'Capsid protein diversity among 'Norwalk-like' viruses.';
RL Virus Genes 20:227-236(2000).

DR EMBL; AJ277614; CAB89095.1; -;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 543 AA; 58216 MW; 11F97A3989CCCS40 CRC64;

Query Match 76.3%; Score 2208; DB 12; Length 543;
Best Local Similarity 74.0%; Pred. No. 2.7e-174;
Matches 404; Conservative 63; Mismatches 75; Indels 4; Gaps 3;

QY 1 MMASKADPQSDAGSAGAGOLVPEVNTADPLMEPVAGPTTAVATAGVNMIDPVIWNNF 60
DB 1 MMASKADPQSDAGSAGAGOLVPEVNTADPLMEPVAGPTTAVATAGVNMIDPVIWNNF 60
QY 61 VQAPQGEFTISPNNTPGDIPLDQLGPHLNPFLSHLSQMYNGWGMVRVILLAGNAFSA 120
DB 61 VQAPQGEFTISPNNTPGDIPLDQLGPHLNPFLSHLSQMYNGWGMVRVILLAGNAFSA 120
QY 121 GKIIICVCPVPGTSSSLTIAQATLFPHVIAVRLTPEIEMPLEDVRNVLVYHTNDNQPTWR 180
DB 121 GKIIICVCPVPGTSSSLTIAQATLFPHVIAVRLTPEIEMPLEDVRNVLVYHTNDNQPTWR 180
QY 181 LVCMLYTPLRTGCGSGNSDSFVAGRVLTAPSDSFLVPPTEIQKTRAFVVPNIPLQ 240
DB 181 LVCMLYTPLRTGCGSGNSDSFVAGRVLTAPSDSFLVPPTEIQKTRAFVVPNIPLQ 240
QY 241 TLNSRPFSLIQMILSPDASQVQFQNGRCLIDGQLGTTTATSGQLFRVRGKINQGAR 300
DB 241 TLNSRPFSLIQMILSPDASQVQFQNGRCLIDGQLGTTTATSGQLFRVRGKINQGAR 300
QY 301 TLNLTVDGKPMAPFSPAPVGPDPFGKCDWHMRISKTPNNTSSGDPMSRSVQTNVQGF 360
DB 301 VLNLTELDCGSPFPAAPAGPDLGSCDWHIEMSKIPNSQNNPIVNSVKPNSQGF 360
QY 361 VPHLSGIQFDEVNHTPDYIGTIEWISQPS-TPGTDINLWEIPDYGSSLSQANLAPPV 419
DB 361 VPHLSGSVTTAID-TAGDTLGTIOWTSQPSNVTVPDVFNTIPOYSSSLAEASQAPV 419
QY 420 FPPGFGALVYFVSAPFPNRRSAPNDVPCLLPQEIYTHFVSEQATMGDAALLHYVDP 479
DB 420 YPPGFNEVYFVWASIPGNSQSPNLVPCLLPQEIYTHFVSEQATMGDAALLHYVDP 479
QY 480 TNRNLGEFKLYPGGYLTCVPSNGVAGPQQLPLNGVFLFVSWSRFYQLKPVGTASTARS 539
DB 480 TNRNLGEFKLYPGGYLTCVPSNGVAGPQQLPLNGVFLFVSWSRFYQLKPVGTASTARS 539
QY 540 LGVRR 545
DB 538 LGVRR 543

RESULT 9
Q91EJ2 ID Q91EJ2 PRELIMINARY; PRT; 544 AA.
AC Q91EJ2
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Nucleocapsid protein.
OS Human calicivirus HU/NLV/Queen's Arms/Leeds/92/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=159480;
RN [1]
RP SEQUENCE FROM N.A.
RA Clegg C.S., Chamberlain J., Green J., Brown D.W.G., Lewis D.;
RT "Analysis of diversity in the capsid and ORF3 proteins of some
RT Norwalk-like viruses circulating in the U.K.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ313030; CAC40987.1; -;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.

[illegible]


```
AC Q918D1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
GN Norwalk-like virus NLV/New Orleans/266/1993/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171834;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/266/1993/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/266/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of Norwalk-like viruses.";
RL J. Infect. Dis. 181:S336-S348 (2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/266/1993/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RA Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414402; AAL12959.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat_1.
SQ SEQUENCE 544 AA; 58354 MW; 50387E99852E801D CRC64;

Query Match 76.0%; Score 2201.5; DB 12; Length 544;
Best Local Similarity 73.9%; Pred. No. 9.5e-174;
Matches 403; Conservative 58; Mismatches 83; Indels 1; Gaps 1;

QY 1 MMASKDAPQSGADGAGAGQVPEVNTADPLPMEPVAGTTAVATAGQVNMIDPVIWNNF 60
DB 1 MMASKDAPQSGADGAGAGQVPEVNTADPLPMEPVAGTTAVATAGQVNMIDPVIWNNF 60
QY 61 VQSPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGNMVRVILLAGNAFSA 120
DB 61 VQSPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGNMVRVILLAGNAFSA 120
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGNMVRVILLAGNAFSA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGNMVRVILLAGNAFSA 120
QY 121 GKIIIVCCVPPGFTSSSLTIAQATLPPHVIADVRLTLEPIEMPLEDVRNVLHYHNDTQPTMR 180
DB 121 GKIIIVCCVPPGFTSSSLTIAQATLPPHVIADVRLTLEPIEMPLEDVRNVLHYHNDTQPTMR 180
QY 181 LVCMLYTLPTRTGGSGNSDSFVAVGRVLTAPSSDFPLVPPTIEQKTRAFVNPILQ 240
DB 181 LVCMLYTLPTRTGGSGNSDSFVAVGRVLTAPSSDFPLVPPTIEQKTRAFVNPILQ 240
QY 181 LLCMLYTLPTRTGGSGNSDSFVAVGRVLTAPSSDFPLVPPTIEQKTRAFVNPILQ 240
DB 181 LLCMLYTLPTRTGGSGNSDSFVAVGRVLTAPSSDFPLVPPTIEQKTRAFVNPILQ 240
QY 241 TILSNRFPESLIQGMILSPDASQVQFNGRCILDGQLGTTTATSGQLFRVKGKINQCAR 300
DB 241 TILSNRFPESLIQGMILSPDASQVQFNGRCILDGQLGTTTATSGQLFRVKGKINQCAR 300
QY 301 TLNLTVDGKPFMAFDSAPVGFPPDFGKCDWMHRIKTPNNTSSGDPNRSVSVQTNVQGF 360
DB 301 TLNLTVDGKPFMAFDSAPVGFPPDFGKCDWMHRIKTPNNTSSGDPNRSVSVQTNVQGF 360
QY 301 ALNLTLDGSPFMAFAPAPAGFPDLGSCDWHIENSKIPINSTQNNPIVNSVKNSQOF 360
DB 301 ALNLTLDGSPFMAFAPAPAGFPDLGSCDWHIENSKIPINSTQNNPIVNSVKNSQOF 360
QY 361 VPHLSIQDFE-VENHPTGDIYGTIEW-----ISQSTPPCTDNLNWEIPDYGSSLSQA 413
DB 361 VPHLSIQDFE-VENHPTGDIYGTIEW-----ISQSTPPCTDNLNWEIPDYGSSLSQA 413
QY 414 ANLAPVPPFPGFEALVYFVSAPFQGNRRNSAPNDVPCLLPQEIYTHFVSEQAPTMDAAL 473
DB 414 ANLAPVPPFPGFEALVYFVSAPFQGNRRNSAPNDVPCLLPQEIYTHFVSEQAPTMDAAL 473
QY 413 SOLAPAVYPPGFEALVYFVSAPFQGNRRNSAPNDVPCLLPQEIYTHFVSEQAPTMDAAL 472
DB 413 SOLAPAVYPPGFEALVYFVSAPFQGNRRNSAPNDVPCLLPQEIYTHFVSEQAPTMDAAL 472
QY 474 LHYVDPDTRNLGEPKLYPGGYLTCPVNGVAGAGPQPLPLNGVFLFVSVWVSRYPQLKPVGT 533
DB 474 LHYVDPDTRNLGEPKLYPGGYLTCPVNGVAGAGPQPLPLNGVFLFVSVWVSRYPQLKPVGT 533
```

Wed Jun 2 09:13:28 2004

```

473 LHVDPDTRNLGEFKLYPGGYLTCVPNSSTGPOQLPDLGDFVFWFASWVSRYQLKPVGT 532
534 ASTARSRLGVR 545
533 AGFARGRLGVR 544

RESULT 13
Q9IV41 PRELIMINARY; PRT; 544 AA.
AC Q8XJ0; 2002 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Valetta/95/Malta.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=122920;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/Valetta/95/Malta;
RX MEDLINE=20404983; PubMed=10949950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236 (2000).
DR EMBL: AJ277616; CAB89097.1; -
DR InterPro: IPR004005; Calici coat.
DR InterPro: IPR008975; Viral cap_coat.
DR Pfam: PF00915; Calici coat; 1.
SQ SEQUENCE 544 AA; 58415 MW; B6741846BD82E6D1 CRC64;

Query Match 75.3%; Score 2179.5; DB 12; Length 544;
Best Local Similarity 73.4%; Pred. No. 6.3e-172;
Matches 400; Conservative 60; Mismatches 84; Indels 1; Gaps 1;

QY 1 MMASKDAPQADGASGAGQLVPEVNTADPLPMEPVAGPTTAVATAGVNMIDPWVNNF 60
DB 1 MMASKDAPQADGAGQLVPEVNTADPLPMEPVAGPTTAVATAGVNMIDPWVNNF 60
QY 61 VQSPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLSQMYNGWGNMVRVILAGNAFSA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLSQMYNGWGNMVRVILAGNAFTA 120
QY 121 GKIIIVCCVPPGFTSSSLTIAQATLPPHVIADVRLTLEPIEMPLEDVRNVLXHTNDQPTMR 180
DB 121 GKVIICCVPPGFTSSSLTIAQATLPPHVIADVRLTLEPIEMPLEDVRNVLXHTNDQPTMR 180
QY 181 LVCMLYTLPLRTGGSGNSDSFVWAGRVLTAPSSDFSLFVPPPTIEQKTRAFVNPILQ 240
DB 181 LLCMLYTLPLRTGGSGNSDSFVWAGRVLTAPSSDFSLFVPPPTIEQKTRAFVNPILQ 240
QY 241 TILSNRFPPLIQGMILSPDASQVQFQNGRCLIDQLLGTTPATSGQLFRVRGKINOGAR 300
DB 241 TILSNRFPPLIQGMILSPDASQVQFQNGRCLIDQLLGTTPATSGQLFRVRGKINOGAR 300
QY 301 TLNLTVDGKPFMAFDSAPVGFDFGKCDHMRISKTPNTSSGDMRSVSVQTNVQGF 360
DB 301 VLNLTKLDGSPFMAFAPAPAGFPDLGSGDHAQVTFQNGRCLTDLGLLGTTPVSAQGLMKRGTSGSK 360
QY 361 VPHLSGIQDFEVNHTGDIYGTIEWISQPTTDLNWEIPDYGSSLSQAANLAPPV 420
DB 361 VPHLSGIQDFEVNHTGDIYGTIEWISQPTTDLNWEIPDYGSSLSQAANLAPPV 420
QY 421 FPPGFEALVYFVSAPFGNNRNSAPNDVPCLLIPQSYITHFVSEQAPTMGDALLHYDDP 480
DB 420 YPLVNEVIVYFMACIRGNQCGSNLVPCLSPQEIYTHFISEQAPIQGEAALLHYDDP 479
QY 481 TNRNLGEFKLYPGGYLTCVPNGVAGAGPQQLPLNGVFLFVSWVSRYQLKPVGTASTARS 540
DB 480 TNRNLGEFKLYPGGYLTCVPNSSTGPOQLPDLGDFVFWFASWVSRYQLKPVGTAGPARGR 539

```

```

QY 541 LGVRR 545
DB 540 LGVRR 544

RESULT 14
Q8XJ0 PRELIMINARY; PRT; 543 AA.
AC Q8XJ0; 2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DE Capsid.
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S2UG1;
RA Katayama K., Kojima S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S2UG1;
RX MEDLINE=22192455; PubMed=1202225;
RA Katayama K., Shirato-Horikoshi H., Kojima S., Kageyama T., Oka T.,
RA Hoshino F.B., Fukushi S., Shinohara M., Uchida K., Suzuki Y.,
RA Gojibori T., Takeda N.;
RT "Phylogenetic Analysis of the Complete Genome of 18 Norwalk-like
RT Viruses.";
RL Virology 299:225-239 (2002).
DR EMBL: AB039774; BAC11813.1; -
DR InterPro: IPR004005; Calici coat.
DR InterPro: IPR008975; Viral cap_coat.
DR Pfam: PF00915; Calici coat; 1.
SQ SEQUENCE 543 AA; 58761 MW; E0DE7489A9C3488C CRC64;

Query Match 75.2%; Score 2178; DB 12; Length 543;
Best Local Similarity 73.3%; Pred. No. 8.4e-172;
Matches 401; Conservative 65; Mismatches 75; Indels 6; Gaps 4;

QY 1 MMASKDAPQADGASGAGQLVPEVNTADPLPMEPVAGPTTAVATAGVNMIDPWVNNF 60
DB 1 MMASKDAPQADGAGQLVPEVNTADPLPMEPVAGPTTAVATAGVNMIDPWVNNF 60
QY 61 VQSPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLSQMYNGWGNMVRVILAGNAFSA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLSQMYNGWGNMVRVILAGNAFTA 120
QY 121 GKIIIVCCVPPGFTSSSLTIAQATLPPHVIADVRLTLEPIEMPLEDVRNVLXHTNDQPTMR 180
DB 121 GKVIICCVPPGFTSSSLTIAQATLPPHVIADVRLTLEPIEMPLEDVRNVLXHTNDQPTMR 180
QY 181 LVCMLYTLPLRTGGSGNSDSFVWAGRVLTAPSSDFSLFVPPPTIEQKTRAFVNPILQ 240
DB 181 LLCMLYTLPLRTGGSGNSDSFVWAGRVLTAPSSDFSLFVPPPTIEQKTRAFVNPILQ 240
QY 241 TILSNRFPPLIQGMILSPDASQVQFQNGRCLIDQLLGTTPATSGQLFRVRGKINOGAR 300
DB 241 TILSNRFPPLIQGMILSPDASQVQFQNGRCLIDQLLGTTPATSGQLFRVRGKINOGAR 300
QY 301 TLNLTVDGKPFMAFDSAPVGFDFGKCDHMRISKTPNTSSGDMRSVSVQTNVQGF 360
DB 301 VLNLTKLDGSPFMAFAPAPAGFPDLGSGDHAQVTFQNGRCLTDLGLLGTTPVSAQGLMKRGTSGSK 360
QY 361 VPHLSGIQDFEVNHTGDIYGTIEWISQPTTDLNWEIPDYGSSLSQAANLAPPV 418
DB 360 VPHLSGSFKEVNV - VAGDYVCTIQTSPSPDPHPNADVDWFWSIDPDYGSNLAESQLAP 418
QY 419 FPPGFEALVYFVSAPFGNNRNSAPNDVPCLLIPQSYITHFVSEQAPTMGDALLHYDD 478
DB 419 VVYPPGFEALVYFVSAPFGNNRNSAPNDVPCLLIPQSYITHFVSEQAPTMGDALLHYDD 478

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 46.7329 Seconds
(without alignments)
3289.030 Million cell updates/sec

Title: US-09-926-799-4
Perfect score: 2891
Sequence: 1 MMASKATPDSADGATGAGQ.....YOLKPVGTAGPARGRLGVRR 544

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2891	100.0	544	AAB49703	Small rou
2	2231.5	77.2	546	AAB49702	Small rou
3	2100	72.6	530	AAR50972	Norwalk v
4	2100	72.6	530	ADC72176	Norwalk v
5	2093	72.4	530	AAB49701	Small rou
6	2054	71.0	530	AAR57091	Small rou
7	1979	68.5	545	AAB49700	Small rou
8	1170.5	40.5	540	AAB49706	Small rou
9	1165.5	40.3	542	AAB49708	Small rou
10	1151.5	39.8	539	AAB49704	Small rou
11	1145	39.6	535	AAB49707	Small rou
12	1143.5	39.6	548	AAB49705	Small rou
13	1134	39.2	548	AAB49709	Small rou
14	1124.5	38.9	550	AAB49710	Small rou
15	970	33.6	541	AAB49710	Small rou
16	309	10.7	579	AAB49704	Small rou
17	268	9.3	568	AAB67462	Amino aci
18	264.5	9.1	547	AAM50108	Feline ca
19	264.5	9.1	671	AAM50107	Feline ca
20	263.5	9.1	668	AAR10686	Feline ca
21	263.5	9.1	668	AAR10686	Feline ca
22	263.5	9.1	669	AAB67461	Amino aci
23	261.5	9.0	623	AAB47044	Feline ca
24	261.5	9.0	623	AAB47043	Feline ca
25	250.5	8.7	622	AAB47045	Feline ca

26	143.5	5.0	2971	7	ADC35084	Adc35084 Human bre
27	143.5	5.0	2972	4	AAB50363	Aab50363 Human SRC
28	143.5	5.0	3118	4	AAB50362	Aab50362 Human SRC
29	142.5	4.9	2735	6	ABR41356	ABr41356 Human DIT
30	141	4.9	2206	2	AAR22210	Aar22210 True type
31	140.5	4.9	2971	3	AAB41231	Aab41231 Human ORF
32	140	4.8	40	5	AAB91273	Aau91273 Norwalk v
33	130	4.5	3930	6	ABU18893	Abu18893 Protein e
34	127	4.4	1127	4	AAB95541	Aab95541 Human pro
35	126.5	4.4	1185	6	ABU18969	Abu18969 Protein e
36	126	4.4	40	5	AAB91274	Aau91274 Norwalk v
37	125.5	4.3	2358	6	ABU17839	Abu17839 Protein e
38	125	4.3	1152	7	ADE59446	Ade59446 Human Pro
39	123.5	4.3	6310	6	ABU39869	Abu39869 Protein e
40	121	4.2	1150	4	AAM40399	Aam40399 Human pol
41	121	4.2	2183	3	AAB03533	Aab03533 Murine fa
42	120.5	4.2	895	3	AAV56804	Aav56804 Clathrin
43	117.5	4.1	1155	2	AAR28047	Aar28047 IRS-1 pro
44	117.5	4.1	2016	4	ABB63911	ABb63911 Drosophill
45	117	4.0	1564	7	ADC86801	Adc86801 Human GPC

ALIGNMENTS

RESULT 1
AAB49703
ID AAB49703 standard; protein; 544 AA.
XX
AC AAB49703;
XX
DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 4.
XX
KW Small round structured virus; SRSV; food poisoning.
XX
OS Small round structured virus.
XX
PN WO200079280-A1.
XX
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000WO-JP004095.
XX
PR 22-JUN-1999; 99JP-00175928.
XX
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
XX
Takeda N, Natōri K, Miyamura T, Kamata K, Sato T, Sato S;
WPI; 2001-080848/09.
DR N-PSDB; AAF29144.
XX
Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.
XX
Claim 1; Page 47-49; 84pp; Japanese.
XX
This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks
SQ Sequence 544 AA;
Query Match 100.0%; Score 2891; DB 4; Length 544;
Best Local Similarity 100.0%; Pred. No. 9.4e-249;

Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MMASKDATPSADGATGAGQGVPEVNTADPIPDPVAGSSTALATAGQVNLIDPWIIINNF 60
Db	1 MMASKDATPSADGATGAGQGVPEVNTADPIPDPVAGSSTALATAGQVNLIDPWIIINNF 60
QY	61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLAGNAFTA 120
Db	61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLAGNAFTA 120
QY	121 GKVIICCVPPGQSRITLSIAQATLFPFHVIAVRTLDPEVPLEDRVNLVYHNDTQPTMR 180
Db	121 GKVIICCVPPGQSRITLSIAQATLFPFHVIAVRTLDPEVPLEDRVNLVYHNDTQPTMR 180
QY	181 LLCMLYTLRTGGASGSDTSFVAVGRVLTCPGDFNFELVPPTEQKTRPFTVPIPLK 240
Db	181 LLCMLYTLRTGGASGSDTSFVAVGRVLTCPGDFNFELVPPTEQKTRPFTVPIPLK 240
QY	241 YLSNSRIPNPIEGMSLSPDOTQNVQFQNGRCTIDGQPLGTTTPVSVSQLCKFRGRTSQOR 300
Db	241 YLSNSRIPNPIEGMSLSPDOTQNVQFQNGRCTIDGQPLGTTTPVSVSQLCKFRGRTSQOR 300
QY	301 VLNLTELDSGSPFMAFAAPAGAPGPDGLGSCDWHIEMSKI PNSTQNNPIVTVNSVKPNSQOF 360
Db	301 VLNLTELDSGSPFMAFAAPAGAPGPDGLGSCDWHIEMSKI PNSTQNNPIVTVNSVKPNSQOF 360
QY	361 VPHLSSTILDENVS-SGGDYIGTQWTSPPSDSGANTFWKIPDYGSSLAESAQLAPAVY 420
Db	361 VPHLSSTILDENVS-SGGDYIGTQWTSPPSDSGANTFWKIPDYGSSLAESAQLAPAVY 420
QY	421 PPGNEVIVVFMAISIPGPNQSGSNLVPCLLPQEIYTHFISEQAPITQGEAALLHYVDDPT 480
Db	421 PPGNEVIVVFMAISIPGPNQSGSNLVPCLLPQEIYTHFISEQAPITQGEAALLHYVDDPT 480
QY	481 NRNLGEPKLYPGGYLTCVPNSSSTGPOQLPLDGVVFASWVSRYQLKPVGTAGPARGRL 540
Db	481 NRNLGEPKLYPGGYLTCVPNSSSTGPOQLPLDGVVFASWVSRYQLKPVGTAGPARGRL 540
QY	541 GVR 544
Db	541 GVR 544
RESULT 2	
ID	AAB49702
AC	AAB49702 standard; protein; 546 AA.
XX	AAB49702;
DT	04-APR-2001 (first entry)
DE	Small round structured virus protein SEQ ID 3.
XX	Small round structured virus; SRSV; food poisoning.
XX	Small round structured virus.
XX	WO200079280-A1.
XX	28-DEC-2000.
XX	22-JUN-2000; 2000WO-JP004095.
XX	22-JUN-1999; 99JP-00175928.
XX	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX	(DENK-) DENKA SEIKEN KK.
XX	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX	WPI; 2001-080848/09.
XX	N-PSDB; AAF29143.
XX	XX

Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.

Claim 1; Page 45-47; 84pp; Japanese.

This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

Sequence 546 AA;

Query Match	77.2%; Score 2231.5; DB 4; Length 546;
Best Local Similarity	74.7%; Pred. No. 6.8e-190;
Matches	407; Conservative 59; Mismatches 78; Indels 1; Gaps 1;
QY	1 MMASKDATPSADGATGAGQGVPEVNTADPIPDPVAGSSTALATAGQVNLIDPWIIINNF 60
Db	1 MMASKDATPSADGATGAGQGVPEVNTADPIPDPVAGSSTALATAGQVNLIDPWIIINNF 60
QY	61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLAGNAFTA 120
Db	61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLAGNAFTA 120
QY	121 GKVIICCVPPGQSRITLSIAQATLFPFHVIAVRTLDPEVPLEDRVNLVYHNDTQPTMR 180
Db	121 GKVIICCVPPGQSRITLSIAQATLFPFHVIAVRTLDPEVPLEDRVNLVYHNDTQPTMR 180
QY	181 LLCMLYTLRTGGASGSDTSFVAVGRVLTCPGDFNFELVPPTEQKTRPFTVPIPLK 240
Db	181 LLCMLYTLRTGGASGSDTSFVAVGRVLTCPGDFNFELVPPTEQKTRPFTVPIPLK 240
QY	241 YLSNSRIPNPIEGMSLSPDOTQNVQFQNGRCTIDGQPLGTTTPVSVSQLCKFRGRTSQOR 300
Db	241 YLSNSRIPNPIEGMSLSPDOTQNVQFQNGRCTIDGQPLGTTTPVSVSQLCKFRGRTSQOR 300
QY	301 VLNLTELDSGSPFMAFAAPAGAPGPDGLGSCDWHIEMSKI PNSTQNNPIVTVNSVKPNSQOF 360
Db	301 VLNLTELDSGSPFMAFAAPAGAPGPDGLGSCDWHIEMSKI PNSTQNNPIVTVNSVKPNSQOF 360
QY	361 VPHLSSTILDENVS-SGGDYIGTQWTSPPSDSGANTFWKIPDYGSSLAESAQLAPAV 419
Db	361 VPHLSSTILDENVS-SGGDYIGTQWTSPPSDSGANTFWKIPDYGSSLAESAQLAPAV 420
QY	420 YPPGNEVIVVFMAISIPGPNQSGSNLVPCLLPQEIYTHFISEQAPITQGEAALLHYVDDPT 479
Db	421 PPGNEVIVVFMAISIPGPNQSGSNLVPCLLPQEIYTHFISEQAPITQGEAALLHYVDDPT 480
QY	480 TNRNLGEPKLYPGGYLTCVPNSSSTGPOQLPLDGVVFASWVSRYQLKPVGTAGPARGR 539
Db	481 TNRNLGEPKLYPGGYLTCVPNGVAGPQQLPLNGVLFVSVWVSRYQLKPVGTASTARSR 540
QY	540 LGVRR 544
Db	541 LGVRR 545

RESULT 3

AAR50972

ID AAR50972 standard; protein; 530 AA.

XX AAR50972;

XX AAR50972;

DT 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 05-OCT-1994 (first entry)

XX Norwalk virus strain 8FIIa protein (encoded by ORF2).

XX

KW Norwalk virus; pathogen; acute gastroenteritis; food poisoning;
 XX seafood contamination; diagnostic assay; calicivirus; small round virus.

OS Norwalk virus; (strain 8FIIa).

XX WO9405700-A2.

XX PD 17-MAR-1994.

XX PF 07-SEP-1993; 93WO-US008447.

XX PR 07-SEP-1992; 92US-00941365.

XX PA (BAYU) BAYLOR COLLEGE MEDICINE.

XX PI Matson DO, Estes MK, Jiang X, Graham DY;

XX DR WPI: 1994-101125/12.

XX DR N-PSDB; AAQ56826.

XX DNA from Norwalk and related viruses - used for preparing prods. for use
 PT in diagnostic assays, detection and vaccines for Norwalk and related
 PT viruses.

XX PS Claim 14; Page 68-70; 156pp; English.

XX CC The Norwalk virus was isolated from stool samples from adult volunteers
 CC infected with safety tested Norwalk virus strain 8FIIa. The coding
 CC sequence is useful for the design of probes for use in diagnostic assays
 CC for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct
 CC PN field.) (Updated on 16-OCT-2003 to standardise OS field)

XX SQ Sequence 530 AA;

Query Match 72.6%; Score 2100; DB 2; Length 530;
 Best Local Similarity 72.7%; Pred. No. 3.5e-178;
 Matches 396; Conservative 54; Mismatches 79; Indels 16; Gaps 8;

QY 1 MMASKDTPSADGATGATGQVPEVNTADPIPDPVAGSSTALATAGQVNLDPWIINF 60
 DB 1 MMASKDTPSADGATGATGQVPEVNTADPIPDPVAGSSTALATAGQVNLDPWIINF 60
 QY 61 VOAPQGETTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMRVRLVLAGNAFTA 120
 DB 61 VOAPQGETTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMRVRLVLAGNAFTA 120
 QY 121 GKVIICVPPCGFQSTLTIAQATLPHVIADVTLDPEVPLEDVRLVLYHND-TQPTM 179
 DB 121 GKVIICVPPCGFQSTLTIAQATLPHVIADVTLDPEVPLEDVRLVLYHND-TQPTM 180
 QY 180 RLCLMLYPLRTGGASGGTDFVWAGRVLTCPDPFNFLLVPPTVEQKTRPFTVPNPL 239
 DB 181 RLVCMLYPLRTGGTGG--DSFVWAGRVWTCPSDPFNFLLVPPTVEQKTRPFTVPNPL 238
 QY 240 KYLSNSRIPNPIEGMSLPDQTONVQFQNGRCITDQPLGTTPFVSQCKPRGRITSGQ 299
 DB 239 SLSNSRAPLPISSGISPDNVQVQFQNGRCITDQPLGTTPFVSQCKPRGRITSGQ 296
 QY 300 RVNLNLTLDGSPFMAFAAPAGFDLGLSDWHIEMSKIPNSSTQNNPIVNSVKNPQQ 359
 DB 297 TVINLTLDGTPFHPFEGPAPIGFDLGLSDWHIEMSKIPNSSTQNNPIVNSVKNPQQ 351
 QY 360 FVPHLSSTLDNVSSGGDIQTIQWTPSPSGGANTNFWKIPDYGSLSAEASOLAPAV 419
 DB 352 FVPHLSSTLDNVSSGGDIQTIQWTPSPSGGANTNFWKIPDYGSLSAEASOLAPAV 409
 QY 420 YPPGFNEVIVYPMASIPGNQSGSNLVPCLLPQBYITHFISEQAPIQCEAALLHYVDPD 479
 DB 410 YPPGFNEVIVYPMASIPGNQSGSNLVPCLLPQBYITHFISEQAPIQCEAALLHYVDPD 465
 QY 480 TNRNLGEFKLYPGGHVLTCPVNSSTGPPQPLDGVFVFNASVSRFYQLKPVGTAGPARGR 539
 DB 466 TNRNLGEFKLYPGGHVLTCPVNSSTGPPQPLDGVFVFNASVSRFYQLKPVGTAGPARGR 525

QY 540 LGVRR 544
 DB 526 LGLRR 530

RESULT 4

ADCT2176
 ID ADC72176 standard; protein; 530 AA.

XX AC ADC72176;

XX DT 18-DEC-2003 (first entry)

XX Norwalk virus protein 2 amino acid sequence.

XX immune response; non-Norwalk virus agent; immunogen; Norwalk virus;
 KW viral pathogen; acute gastroenteritis; virucidal; antiviral vaccine.

XX OS Norwalk virus.

XX PN US6572862-B1.

XX PD 03-JUN-2003.

XX PF 07-JUN-1995; 95US-00486049.

XX PR 08-NOV-1989; 89US-004333492.

XX PR 27-APR-1990; 90US-00515993.

XX PR 27-AUG-1990; 90US-00573509.

XX PR 06-MAY-1991; 91US-00696454.

XX PA (BAYU) BAYLOR COLLEGE MEDICINE.

XX PI Estes MK, Jiang X, Graham DY;

XX DR WPI: 2003-776005/73.

XX DR N-PSDB; ADC72174.

XX Inducing an immune response against non-Norwalk virus agents, comprises
 PT administering an immunogen recombinantly expressed from a cDNA from
 PT Norwalk virus.

XX PS Example 4; SEQ ID NO 3; 45pp; English.

XX CC This invention relates to a novel method of inducing an immune response
 CC in an individual against Norwalk virus and non-Norwalk virus agents, by
 CC orally or parenterally administering an immunogen recombinantly expressed
 CC or synthesised from a cDNA of Norwalk virus given in the specification.
 CC Norwalk virus is one of the most important viral pathogens, causing acute
 CC gastroenteritis. The invention may be used for the development of
 CC compounds with virucidal activity or an antiviral vaccine. The present
 CC sequence is the amino acid sequence of a protein encoded by the Norwalk
 CC virus genome of the invention.

XX SQ Sequence 530 AA;

Query Match 72.6%; Score 2100; DB 7; Length 530;
 Best Local Similarity 72.7%; Pred. No. 3.5e-178;
 Matches 396; Conservative 54; Mismatches 79; Indels 16; Gaps 8;

QY 1 MMASKDTPSADGATGATGQVPEVNTADPIPDPVAGSSTALATAGQVNLDPWIINF 60
 DB 1 MMASKDTPSADGATGATGQVPEVNTADPIPDPVAGSSTALATAGQVNLDPWIINF 60
 QY 61 VOAPQGETTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMRVRLVLAGNAFTA 120
 DB 61 VOAPQGETTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMRVRLVLAGNAFTA 120
 QY 121 GKVIICVPPCGFQSTLTIAQATLPHVIADVTLDPEVPLEDVRLVLYHND-TQPTM 179
 DB 121 GKVIICVPPCGFQSTLTIAQATLPHVIADVTLDPEVPLEDVRLVLYHND-TQPTM 180

180 RLLCMLYPLRTGASGSDTSFVAGRVLTCPGPDENFLVPTVEQKTRPFTVPNIPL 239
 181 RLVCMYPLRTGGGTG--DSFVAGRVMTCPSPDENFLVPTVEQKTRPFTLPNPL 238
 240 KYLSNRIPNPTLGGSLSDPDTQNVQFQNGRGTIDGQPLGTPVSVSOLCKFRGRTSQ 299
 239 SLSNSRAPLPLISSMGISPDNVQSVQFQNGRGTLDGRLVGTTPVSLSHVAKIRG--TSNG 296
 300 RVNLNLTDLGSPFMAAPAPAGFPDLGSCDWHIEMSKIPIKNSSTQNNPIVTVNSVKPNSQ 359
 297 TVINLTDLGTPFHPFEGPAPIGFPDLGSCDWHIEMSKIPIKNSSTQNNPIVTVNSVKPNSQ 351
 360 FVPHLSITLDENVSSGGYICTIOWTSPDSGGANTFVKIPDYGSSLAESAOLAPAV 419
 352 FVPHLGSII--QANGIGSGNYGVLSWSPSPHSGSQVDLWKIPNYGSSITEATHLAPSV 409
 420 YPGFNEVIVFMASIPGPNQSGSNLVPCLLPQEIYTHIFISEQAPIQGEAALLHYVDP 479
 410 YPGFGEVLVFFMSKMPG--GAYNL-PCLLPQEIYSHLASEQAPTVEAALLHYVDP 465
 480 TNRNLGEPKLYPGCYLTCVPNSSTGPOQLPLDGVFVAFASWVSFRFVQLKPVGTAGPARGR 539
 466 TGRNLGEPKAYPDGFLTCVPNGASSGPOQLPINGVFVSVWSRFFVQLKPVGTASSARGR 525
 540 LGVRR 544
 526 LGLRR 530

RESULT 5
 AAB49701
 ID AAB49701 standard; protein; 530 AA.
 XX
 AC AAB49701;
 DT 04-APR-2001 (first entry)
 XX
 DE Small round structured virus protein SEQ ID 2.
 XX
 KW Small round structured virus; SRSV; food poisoning.
 XX
 OS Small round structured virus.
 XX
 PN WO200079280-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 22-JUN-2000; 2000WO-JP004095.
 XX
 PR 22-JUN-1999; 99JP-00175928.
 XX
 PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 XX
 PI (DENK-) DENKA SEIKEN KK.
 XX
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 XX
 DR WPI; 2001-080848/09.
 XX
 DR N-PSDB; AAF29142.
 XX
 PT Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX
 PS Claim 1; Page 42-45; 84pp; Japanese.
 XX
 CC This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAB20141 -
 CC AAB20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks

Sequence 530 AA;
 Query Match 72.4%; Score 2093; DB 4; Length 530;
 Best Local Similarity 72.7%; Pred. No. 1.5e-177; Indels 16; Gaps 8;
 Matches 396; Conservative 53; Mismatches 80;

1 MMASKDATPSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGOVNIIDWIINNPF 60
 1 MMASKDATPSVDGASGAGQLVPEVNASDPLAMPVAGSSTAVATAGOVNIDWIINNPF 60
 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHNLFSLHSONYNGWGNMRVVRVVLAGNAFTA 120
 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHNLFSLHSONYNGWGNMRVVRVVLAGNAFTA 120
 121 GKVIICVPPGQSRRTLSIAQATLPHVIAVDTLPDVEPLEDVRNLYNNND-TQPTM 179
 121 GKIIVSCIPPGGSHNLIQAATLPHVIAVDTLPDVEPLEDVRNLYNNND-TQPTM 180
 180 RLLCMLYPLRTGGASGSDTSFVAGRVLTCPGPDENFLVPTVEQKTRPFTVPNIPL 239
 181 RLVCMYPLRTGGGTG--DSFVAGRVMTCPSPDENFLVPTVEQKTRPFTLPNPL 238
 240 KYLSNRIPNPTLGGSLSDPDTQNVQFQNGRGTIDGQPLGTPVSVSOLCKFRGRTSQ 299
 239 SLSNSRAPLPLISSMGISPDNVQSVQFQNGRGTLDGRLVGTTPVSLSHVAKIRG--TSNG 296
 300 RVNLNLTDLGSPFMAAPAPAGFPDLGSCDWHIEMSKIPIKNSSTQNNPIVTVNSVKPNSQ 359
 297 TVINLTDLGTPFHPFEGPAPIGFPDLGSCDWHIEMSKIPIKNSSTQNNPIVTVNSVKPNSQ 351
 360 FVPHLSITLDENVSSGGYICTIOWTSPDSGGANTFVKIPDYGSSLAESAOLAPAV 419
 352 FVPHLGSII--QANGIGSGNYGVLSWSPSPHSGSQVDLWKIPNYGSSITEATHLAPSV 409
 420 YPGFNEVIVFMASIPGPNQSGSNLVPCLLPQEIYTHIFISEQAPIQGEAALLHYVDP 479
 410 YPGFGEVLVFFMSKIPG--GAYSL-PCLLPQEIYSHLASEQAPTVEAALLHYVDP 465
 480 TNRNLGEPKLYPGCYLTCVPNSSTGPOQLPLDGVFVAFASWVSFRFVQLKPVGTAGPARGR 539
 466 TGRNLGEPKAYPDGFLTCVPNGASSGPOQLPINGVFVSVWSRFFVQLKPVGTASSARGR 525
 540 LGVRR 544
 526 LGLRR 530

RESULT 6
 AAR57091
 ID AAR57091 standard; protein; 530 AA.
 XX
 AC AAR57091;
 XX
 DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 05-OCT-1994 (first entry)
 XX
 DE Small round virus SRSV/KY/89 capsid protein.
 XX
 KW pathogen; acute gastroenteritis; food poisoning; seafood contamination;
 KW diagnostic assay; human calcivirus; small round virus; SRSV; KY89;
 KW Norwalk virus; capsid protein.
 XX
 OS Small round structured virus.
 XX
 PN WO9405700-A2.
 XX
 PD 17-MAR-1994.
 XX
 PF 07-SEP-1993; 93WO-US008447.
 XX
 PR 07-SEP-1992; 92US-00941365.
 XX


```
PA (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI Matson DO, Bates MK, Jiang X, Graham DY;
XX
XX WPI; 1994-101125/12.
XX DR N-PSDB; AAQ56832.
XX
XX DNA from Norwalk and related viruses - used for preparing prods. for use
PT in diagnostic assays, detection and vaccines for Norwalk and related
PT viruses.
XX
XX Example 7; Fig 13a; 156pp; English.
XX
XX The known sequence for Norwalk virus was used to obtain the sequence of
XX other Norwalk-related viruses such as SRSV/KI/89, an agent from a stool
XX from an outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide
XX cDNA sequence includes part of the polymerase region and the capsid
XX region of the genome; the deduced amino acid sequences are AAR57092 and
XX AAR57091, respectively. Expression of fragments and derive. Of Norwalk-
XX related viruses permits development of diagnostic assays to detect
XX antibodies, antigens, viral genetic material or antivirals. (Updated on
XX 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS
XX field.)
XX
XX Sequence 530 AA;
XX
XX Query Match 71.0%; Score 2054; DB 2; Length 530;
XX Best Local Similarity 71.4%; Pred. No. 4.5e-174;
XX Matches 389; Conservative 55; Mismatches 85; Indels 16; Gaps 8;
XX
XX 1 MMASKDATPSADGATGAGQLVPEVNTADPIDPVAGSSTALATAGQVNLDPWIINF 60
XX 1 MMASKDATSSVDGASQVLVPEVNASDPLAMDPAVAGSSTAVATAGQVNPIDWIINF 60
XX
XX 61 VOAPQGETTISPNTTGGVLDLQGLPHLNPPLSHLSOMYNGWGMVRVVLGNAFTA 120
XX
XX 61 VOAPQGETTISPNTTGGVLDLQGLPHLNPPLSHLSOMYNGWGMVRVVLGNAFTA 120
XX
XX 121 GKVIICCVPPGQSRSTLSIAQATLPPHVIADVRTLDPEVEPLEDVRLVYHNND-TQPTM 179
XX 121 GKIIIVSCIPPFGSGSQQLITIAQATLPPHVIADVRTLDPEVEPLEDVRLVYHNNDNQTM 180
XX
XX 180 RLLCMLYTLRTGGASGGTDSFVAVAGRLTCRGPDPFNLFLVPPVTEOKTRPFTVNPIL 239
XX 181 RLVCMLYTLPLSTGGGTG--DSFVAVAGRLTCRGPDPFNLFLVPPVTEOKTRPFTVNPIL 238
XX
XX 240 KYLSNSRIPNPIEGMSLSPDQNTQVQFONGRCTIDGQPLGTTTPVSVQLCCKRGRITSGQ 299
XX 239 SLSNSRAPLPIPSGMIISPDNVQSVQFONGRCTLDGRLVGTTPVSLSHVAKIRG--TSNG 296
XX
XX 300 RVLNLTLDGSPFMAFAAPAGFPDLGSCDWHIEMSKIPNSSTQNNPIVNSVKPNSQQ 359
XX 297 TVINLTLDGTPFHPPEGPAPIGFDLGGCDWHIEMSKIPNSSTQNNPIVNSVKPNSQQ 352
XX
XX 360 FVPHLSSTLDBNSVSGGDIYIGTIQWTSPPDSGGANTNFWKIPDYGSSLAESQALAPAV 419
XX 353 -VPHLGS1--QANGIGSGNYIGVLSWVSPSPSPSGVQDLWKIPYGSSTIETHALPSV 409
XX
XX 420 YPPGNEVIVPMASIPGNSQGSNVLVPCLLPQBYIITHFISEQAPIQGEAALLHYVDPD 479
XX 410 YSPGEGEVLVFMKIPGP--GGDSL-PCLLPQBYIISHLASEQAPTGVGEGLLHYVDPD 465
XX
XX 480 TNRNLGEFKLYPGGYLTCPVNSSSTGPPQPLDGVFVFASVSRFYQLKPVGTAGPARGR 539
XX 466 TDRNLGEFKAYPDGFLTCVNGASSGPQQLPVGUVFVSVSRFYQLKPVGTASTARGR 525
XX
XX 540 LGVRR 544
XX 526 LGLRR 530
XX
XX RESULT 7
XX AAB49700
```

```
ID AAB49700 standard; protein; 545 AA.
XX
XX AC AAB49700;
XX
XX DT 04-APR-2001 (first entry)
XX
XX DE Small round structured virus protein SEQ ID 1.
XX
XX KW Small round structured virus; SRSV; food poisoning.
XX
XX OS Small round structured virus.
XX
XX PN WO200079280-A1.
XX
XX PD 28-DEC-2000.
XX
XX PF 22-JUN-2000; 2000WO-JP004095.
XX
XX PR 22-JUN-1999; 99JP-00175928.
XX
XX PA (NTNA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX
XX PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX DR WPI; 2001-080848/09.
XX DR N-PSDB; AAF29141.
XX
XX PT Kit for the detection and typing of small round-structured virus (SRSV)
XX PT strains for investigation of food poisoning outbreaks, contains
XX PT antibodies.
XX
XX PS Claim 1; Page 40-42; 84pp; Japanese.
XX
XX CC This invention relates to a kit for the detection and typing of small
XX CC round structured virus (SRSV) strains. The kit contains antibodies
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX CC used for detecting and typing strains of SRSV in order to prevent the
XX CC spread of infection and to examine the epidemiology of outbreaks
XX
XX SQ Sequence 545 AA;
XX
XX Query Match 68.5%; Score 1979; DB 4; Length 545;
XX Best Local Similarity 67.0%; Pred. No. 2.3e-167;
XX Matches 368; Conservative 70; Mismatches 101; Indels 10; Gaps 6;
XX
XX 1 MMASKDATPSADGATGAGQLVPEVNTADPIDPVAGSSTALATAGQVNLDPWIINF 60
XX 1 MMASKDATPNTMDGTSGAGQLVPEANTAEPISTMEPVAGATAAATAGQVNMIDPWIMNY 60
XX
XX 61 VOAPQGETTISPNTTGGVLDLQGLPHLNPPLSHLSOMYNGWGMVRVVLGNAFTA 120
XX 61 VOAPQGETTISPNTTGGVLDLQGLPHLNPPLSHLSOMYNGWGMVRVVLGNAFTA 120
XX
XX 121 GKVIICCVPPGQSRSTLSIAQATLPPHVIADVRTLDPEVEPLEDVRLVYHNNDQPTMR 180
XX 121 GKIIIVSCIPPFGSGSQQLITIAQATLPPHVIADVRTLDPEVEPLEDVRLVYHNNDQPTMR 180
XX
XX 181 LLLCMLYTLRTGGASGGTDSFVAVAGRLTCRGPDPFNLFLVPPVTEOKTRPFTVNPIL 240
XX 181 LVCMLYTLRAGSSSGTDPPIAGRLVTCRGPDPFNLFLVPPVTEOKTRPFTVNPIL 240
XX
XX 241 YLSNSRIPNPIEGMSLSPDQNTQVQFONGRCTIDGQPLGTTTPVSVQLCCKRGR1--TSG 298
XX 241 TLSNSRVPESLTKSMWVSRDHQGMVQFONGRVTLDGQLGQTTPTPSAQLCCKIRGSVFHANG 300
XX
XX 299 QRVNLNLTLDGSPFMAFAAPAGFPDLGSCDWHIEMSKIPNSSTQNNPIVNSVKPNSQ 358
XX 301 GNGYNLTLDGSPHAFESPAPIGFPDLGECDDWHEAS--PTTQNTGDVLIKQINVKQES 358
XX
XX 359 QFVPHLSSITLD--ENVSSGGDIYIGTIQWTSPPDSGGANTNFWKIPDYGSSLAESQAL 416
```

359 AFAPHLGTOAGLSVSVNTNMIAGLWVSVSDGHRGDDVPWVIRYSGSTLTERAQLA 418
 417 PAVYPPGNEVIVFMASIPGPNQSGSNL-VPCLLPQXYITHFISQAPIQGBAALLHY 475
 419 PRIYPFGGEALVFVMSDFPIAH--GTNGLSVPCITPQEFVTHFVNEQAPTRGEAALLHY 476
 476 VDPDTRNLNIGPKLYPGGYLTCVPSNSTGTGQQLPLDGVFVFSWVSRYOLKPVGTAGP 535
 477 LDPDTHRLNIGPKLYPEGFMTCPVNSSTGTGPTLPINGVVFVSWVSRYOLKPVGTAGP 536
 536 ARGRGLGVR 544
 537 A-CRUGIR 544
 RESULT 8
 AAB49706
 ID AAB49706 standard; protein; 540 AA.
 AC AAB49706;
 XX
 DT 04-APR-2001 (first entry)
 DE Small round structured virus protein SEQ ID 7.
 DE Small round structured virus; SRSV; food poisoning.
 KW Small round structured virus.
 OS Small round structured virus.
 PN WO200079280-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 22-JUN-2000; 2000WO-JP004095.
 XX
 PR 22-JUN-1999; 99JP-00175928.
 XX
 PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.
 XX
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 XX
 PI WPI; 2001-080848/09.
 DR N-PSDB; AAF29147.
 XX
 PT Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX
 PS Claim 1; Page 54-57; 84pp; Japanese.
 XX
 CC This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX
 SQ Sequence 540 AA;
 Query Match 40.5%; Score 1170.5; DB 4; Length 540;
 Best Local Similarity 45.8%; Pred. No. 3.2e-95;
 Matches 258; Conservative 85; Mismatches 170; Indels 53; Gaps 16;
 1 MMMAKDATPSADGAGQVPEVNTADPIPIIDPVAGSSTALATAGVNLIDPWNNF 60
 1 MKMASNDATPSNDGAAG--LVPESSNN-EMAALEPVVGASLAAPVTGTNIIIDPWIRNF 56
 61 VQAPQGEFTISPNNTGCVLFDLQGLPHLNPLSLHLSOMYNGVWGNMVRVVLGNAGNFTA 120
 57 VQAPNGEFTVSPRNSPGEILLVNLDELPELNPYLAHLARMYNGVAGGMEVQVMLAGNFTA 116

121 GKVIICVPPGQSRSLIAQATLFPBHVADRTLDPEVEPLEDVRNVLVH--NNDTQPTM 179
 117 GKIIIAAAPPYPPVPPVNLSPSQITMPPHHVILIDVRILEPVLPMPPDVRSTLPHFNQKDEPKM 176
 180 RLICMLYTPLRGTGGAGGTDSEFWVAGRVLTCPGDFNPLFLVPPPTVEOKTRPFTVPIPL 239
 177 RIVAMLYTPLRNGS--GDDVFTVSCRILTRPSPEFDFLYLVPPTVESKTKPFTLPVLT 234
 240 KYLSNRIPNPIEGMSLSPDQTONVQFONGRGTIDGQPLGTTTPVSVSOLCKFRGRIT--- 296
 235 GELNSRFPPLSIDEMVTSFNESIVVQPNQGRVTLDELGLLGTTLQACNICIRKVKTCOV 294
 297 -SGORVNL--TELDGSPF-MAFAAPAPAGFPDLGSCDWHIEMSKIPNSQNNPIVINS 352
 295 PSEQHWNLEITNLNGTQDPTDDVPALGVDFPAGEVGVLSQRNRSNPANRAHDAV 354
 353 VKPNSQQFVPHLSITLDENVSSGGDIYIGTIOWTSPSPSGGANTNF----- 399
 355 VATYSDKYTPKGLV-----QIGT--WNT--NDVENOPTKFTPIGLNEVANGHR 399
 400 ---WKIPDYGSSLAASQLAPAVP--PGFNEVIVVFMASIPGPNQSGSPNLVPCLLPOE 454
 400 FEQWTLPRYSGALTNNNLAPAVAPLPG--ERLLFFRSYVPLKGGFGNP-AIDCSVPQE 456
 455 YITHFISEQAPIQGEAALLHYVDPDTRNLNIGPKLYPGGYLTCVPSNSTGTGQQLPLDGV 514
 457 WVQHIFYQESAPSLGDVALVRYVNPDTGRVLFEAKLHKGGLTV--SSTSTGPPVVVPANGY 514
 515 FVFASWVSRYOLKPVGTAGPARGL 540
 515 FKFDWYNQFYSLAPMGT--GNRRRV 539
 RESULT 9
 AAB49708
 ID AAB49708 standard; protein; 542 AA.
 XX
 AC AAB49708;
 XX
 DT 04-APR-2001 (first entry)
 DE Small round structured virus protein SEQ ID 9.
 DE Small round structured virus; SRSV; food poisoning.
 KW Small round structured virus.
 OS Small round structured virus.
 PN WO200079280-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 22-JUN-2000; 2000WO-JP004095.
 XX
 PR 22-JUN-1999; 99JP-00175928.
 XX
 PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.
 XX
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 XX
 PI WPI; 2001-080848/09.
 DR N-PSDB; AAF29149.
 XX
 PT Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX
 PS Claim 1; Page 59-61; 84pp; Japanese.
 XX
 CC This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC

CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX
 XX Sequence 542 AA;

Query Match 40.3%; Score 1165.5; DB 4; Length 542;
 Best Local Similarity 45.3%; Pred. No. 9.1e-95;
 Matches 256; Conservative 89; Mismatches 171; Indels 49; Gaps 16;
 QY 1 MMASKDATPSADGATGAGOLVPE-VNTADPTIPDPVAGSSTALATAGOVNLIDPWINN 59
 Db 1 MMASNDAAPSND---GAASLVPEGIN--ETWLEFVAGASIAAPVAGQTNIIDPWIRTN 55
 QY 60 FVQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGAGNFT 119
 Db 56 FVQAPNGEFTVSPRSPGELLNLNLELGPDLNPLAHLSRMYNGAGGVEVQVLLAGNFT 115
 QY 120 AGKVIICVPPGQFQSGRTLSIAQATLPPHVIADVRTLDPEVPLEDVRNVLVYH--NNDTQPT 178
 Db 116 AGKILFAAIPPNFLVDMISPAQITMLPHLIVDVRTLEPITMLPDVRNVVYHFNQPPQR 175
 QY 179 MELLCLWLYPLRTGCGASGCTDSFVAGRVLTCPGDPENFLVAPTVVEQKTRPFTVNP 238
 Db 176 MELVAMLYPLRSNGS--GDDVFTVSCRVLTRPTDPDFEYLLVPPSVSKTKPFTLPILT 233
 QY 239 LKYLNSRIPNPIEGMSLSPDOTQNVQFQNGRCTIDGQPLGTPVSVSOLCKFRGRITS- 297
 Db 234 ISELNSRFPPIEQLYTAPNETNVVQCGRCTLDGELQGTQLLSAVCFLOQRTVAD 293
 QY 298 -----GQRVNLNTELDGSPF--MAFAAPAGPFDLGSCDWHIEMSKIIPNSSTQNNPIVTN 351
 Db 294 NGDNWDONLLQTYENGASYDPTDEVPAPLGTDQFSGMLYGV-----LTQDNVNVST 345
 QY 352 S-----VKPNSQOFVPHLSITLDENSVSGDYIGTIQWT--SPPSDGGANTNF- 399
 Db 346 GEAKNAKGIYISTTSKGTFFKIGSIGL-----HSITEHVHPNQOSRFTPVGVAVENTPFG 401
 QY 400 -WKIPDYGSSLAESAOLAPAVYPPGFNEVIVYFMASIP--GPNQSGSPNLVPCLLPQEYIT 457
 Db 402 QWVLYHAGSLALNTNLAPAVAPTFFGQLLFRFRSRVPCVQGLQODAFIDCLLQEWVN 461
 QY 458 HFISBOAPIQGEAALLHYVDPTNRLNGELFKLYPGGYLTCTVPNSSTG--POOLPLDGVF 515
 Db 462 HFYQEAAPSQADVALIRYVNPDTGRTLFEAKLHRSGFITV-----SHTGAYPLVWPPNGHF 517
 QY 516 VFASVSVRYQLKPVGTAGPARGRL 540
 Db 518 RPDWVNVQFYSIAPMGT--GNGRRRI 541

RESULT 10
 AAB49704
 ID AAB49704 standard; protein; 539 AA.

XX AAB49704;

XX 04-APR-2001 (first entry)

DE Small round structured virus protein SEQ ID 5.

XX Small round structured virus; SRSV; food poisoning.

XX Small round structured virus.

XX WO200079280-A1.

XX 28-DEC-2000.

XX 22-JUN-2000; 2000WO-JP004095.

XX 22-JUN-1999; 99JP-00175928.

XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.
 XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 XX WPI: 2001-080848/09.
 DR N-PSDB; AAF29145.
 XX Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX Claim 1; Page 50-52; 84pp; Japanese.
 PS This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX
 SQ Sequence 539 AA;

Query Match 39.8%; Score 1151.5; DB 4; Length 539;
 Best Local Similarity 44.6%; Pred. No. 1.6e-93;
 Matches 254; Conservative 80; Mismatches 175; Indels 61; Gaps 16;

QY 1 MMASKDATPSADGATGAGOLVPE-VNTADPTIPDPVAGSSTALATAGOVNLIDPWINN 60
 Db 1 MMASNDANPS--DGST--ANLVPEVNN--EVMALPEVGAATAAPVAGQNVDPWIRNNF 56
 QY 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGAGNFT 120
 Db 57 VQAPQGEFTVSPRNPAGEILMSAPLGPDLNPLAHLSRMYNGAGGVEVQVLLAGNFT 116
 QY 121 GKVIICVPPGQFQSGRTLSIAQATLPPHVIADVRTLDPEVPLEDVRNVLVYH--NNDTQPTM 179
 Db 117 GKIIIPAAVPPNFPFTBGLSPSQVTWPFPHIIVDVQRLEPVLILPLDVRNPFYHYNQSNDS 176
 QY 180 RLICMLYPLRTGCGASGCTDSFVAGRVLTCPGDPENFLVAPTVVEQKTRPFTVNP 239
 Db 177 KLIALMYTPLRANNA--GDDVFTVSCRVLTRPSDFDFILVPPFTVESRTKFTTVPILT 234
 QY 240 KYLSNSRIPNPIEGMSLSPDOTQNVQFQNGRCTIDGQPLGTPVSVSOLCKFRGRIT--- 296
 Db 235 EEMSNSRFPPIEKLITYGPFSSAFVVPQNGRCTTDDVLLGTTQLSAVNICTRGVDTHIA 294
 QY 297 -SGQRVNLNTELDGSPF--MAFAAPAGPFD--LGSCDWHIEMSKIIPNSSTQNNPIVTNSV 353
 Db 295 GSHDYTMNLASQWNSYDPTBEEIPAPLGTDPFVGKIQGLMGTQTTREDGSTRAHKA---TV 351
 QY 354 KPNSSQOFVPHLSITLDENSVSGDYIGTIQWTSPSPD--SGGANTNF----- 399
 Db 352 STGSVHFTPKL-----GSVQYTTDNTNDFTQNTKFTTPGVIGDGNH 395
 QY 400 -----WKIPDYGSSLAESAOLAPAVYPPGFNEVIVYFMASIPGPNQSGSPNL--VPC 452
 Db 396 QNEPQQWVLPNYSGRTHNVHLAPAVAPTFFGQLLFRSTMPG--CSGYPMNNDCLLP 453
 QY 453 QEYITHFISEQAPIQGEAALLHYVDPTNRLNGELFKLYPGGYLTCTVPNSSTGPOOL--P 510
 Db 454 QEWQHFCQEAAPASQDVALLRFVNPDTGRTLFECKLHKSQVTV---AHTGPHDLVIP 509
 QY 511 LDGTVFVFSVSVRYQLKPVGTAGPARGRL 540
 Db 510 PNGYFRFDSWVNVQFYTLAPMGNGAGRRRL 539

RESULT 11
 AAB49707
 ID AAB49707 standard; protein; 535 AA.

us-09-926-799-4.rag

Wed Jun 2 09:13:28 2004

```

XX
AC
XX
XX
DT
XX
DE
XX
KW
XX
OS
XX
XX
PN
XX
PD
XX
PF
XX
PR
XX
XX
PA
XX
PI
XX
PI
XX
DR
XX
XX
PT
XX
PT
XX
XX
PS
XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
XX
XX
SQ
Sequence 535 AA;

Query Match 39.6%; Score 1145; DB 4; Length 535;
Best Local Similarity 45.1%; Pred. No. 66-93;
Matches 249; Conservative 95; Mismatches 178; Indels 30; Gaps 14;

QY 1 MMASKDATPSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGQVNLIDPWIIINF 60
DB 1 MKMASNDAPSDGAG--LVPEANN-ETWALEPVAGASTAAPLTGQNNIIDPWIRLN 56
QY 61 VQAPQGEFTISPNTPGVDLQGLPHNPFSLHLSOMYNGWGNMVRVVLGNAPTA 120
DB 57 VQAPNGEFTVSPRNSPGEVLLNLELGPENLYLAHLRMYNGYAGGVEVQVLLAGNAPTA 116
QY 121 GKVIICCVPPGQSRTLSAQTALPHVITADVRLDPEVLEDRVNLVXH--NNDOTPTM 179
DB 117 GKLVAFAVPPHPFLENISPGQITMFPVHVIDVRLLEPVLPLPDVRRNFFHYNQNEPRM 176
QY 180 RLCLMLYTLRTGGAGSGTDSFVAGRVLTCPGDFNLFVLPVPTVEOKTRPFTVPIPL 239
DB 177 RLVLAMLYTLRSGS--GDDVFTVSCRVLTRSPDPDFNLYLPPTLESKTPFTPLITI 234
QY 240 KYLSNRIENPIEGMSLSDQNTQVQFONGRGTIDQPLGTPVSVSOLCKPRGRI--- 296
DB 235 GELTNSRFPVPIDELITSNPSLVQVQNGRCALDGLQGTQLLPTALCSFRGRINQKV 294
QY 297 SQQR---VLNLTLDGSPMAFA-APAPAGFPDGLGCDWHIEMSKIPNSSTQNNPIVNS 352
DB 295 SGENHWNQVMTINGTPDPTGDVPAPLGTDFSGKLFGLVLSQRDHDNACRSHDAV--- 351
QY 353 VKPNSQOFVPHLSITL-----DENVSSGGDYGITQWTSPPSDSGGANTFNKIPDYGS 408

```

352 IATNSAKFTPKLGAIQITWEEDDV-----HINQPTKFTPVGLFENEGFNQWTLPNYSGA 406

409 LAEASQAPAYPPGFNEVIVFVNASIPGPNQSGPNLVPCLLPQEVITHFISEQAPIQG 468

407 LTLNMGAPPVAPTTPGQILFFRSHIPLKGGVADP--VIDCLLPQEWIQHLYQESAPSQS 465

469 EAALLHYVDPTNRIIGFKLYPGCYLTCVPSNSSSTGPOQLPLDGVFVFAVSWRFRYOLK 528

466 DVALIRFTNPDIGRVLFEAKLHRSYIT-VANTGSR-PIVVPANGYFRFDTWVNFYSLA 523

529 PVGTAGPARGRL 540

524 PMGT-GNGRRRV 534

RESULT 12

AAB49705

ID AAB49705 standard; protein; 548 AA.

XX

AC AAB49705;

XX

DT 04-APR-2001 (first entry)

XX

DE Small round structured virus protein SEQ ID 6.

XX

KW Small round structured virus; SRSV; food poisoning.

XX

OS Small round structured virus.

XX

PN WO200079280-A1.

XX

PD 28-DEC-2000.

XX

PF 22-JUN-2000; 2000WO-JP004095.

XX

PR 22-JUN-1999; 99JP-00175928.

XX

PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

PA (DENK-) DENKA SEIKEN KK.

XX

PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

XX

XX WPI; 2001-080848/09.

DR N-PSDB; AAF29146.

XX

PT Kit for the detection and typing of small round-structured virus (SRSV)

PT strains for investigation of food poisoning outbreaks, contains

PT antibodies.

XX

PS Claim 1; Page 52-54; 84pp; Japanese.

XX

XX This invention relates to a kit for the detection and typing of small

CC round structured virus (SRSV) strains. The kit contains antibodies

CC directed against peptides represented in sequences AAB49700 - AAB49710,

CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -

CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is

CC used for detecting and typing strains of SRSV in order to prevent the

CC spread of infection and to examine the epidemiology of outbreaks

XX

SQ Sequence 548 AA;

Query Match 39.6%; Score 1143.5; DB 4; Length 548;

Best Local Similarity 44.5%; Pred. No. 8.5e-93;

Matches 252; Conservative 86; Mismatches 183; Indels 45; Gaps 15;

QY 1 MMASKDATPSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGQVNLIDPWIIINF 60

DB 1 MKMASNDAPSDGAG--LVPEINN-EMALDPPVAGAAIAAPLTGQNNIIDPWIMNF 56

QY 61 VQAPQGEFTISPNTPGVDLQGLPHNPFSLHLSOMYNGWGNMVRVVLGNAPTA 120

DB 57 VQAPGGEFTVSPRNSPGEVLLNLELGPENLYLAHLRMYNGYAGGVEVQVLLAGNAPTA 116

QY 121 GKVIICUPPGFQSRITSLAQATLPHVIAVRLTDPVEVPLEDVRLVLYH--NNDTQPTM 179
Db 117 GKIIIFAAIPPNPIDNLSAAQITMCPHIVDVRLQEPVNLMPDVRRNFFHYNQSDSRL 176
QY 180 RLLCMLYPLRTGGASGGTDSFVAVAGVLTCPGDPNLFVLPPTVEOKTRPFTVPIPL 239
Db 177 RLIAMLYPLRLAN--NSGDDVFTVSCRVLTRPSDFSNFLVPPTVESKTKFTLPIITI 234
QY 240 KYLSNSRIPNPIEGMSLSPDOTQNVQFQNGRCTIDGQPLGTTTPVSVSOLCKFRGRI--- 295
Db 235 SEMSNRPFPVPIESLHTSPTENIVVQCQNGRVTLDELMTGTTQLLPSQICAPRGVLTST 294
QY 296 -----TSGORVLN-----LTLDGSPF--MAFAAPAPAGFPDLGSCDWHIEMSKIPN 340
Db 295 SRASDAQATATPRLFNYYHVLQDLNLTGTPDPAEDI FGLGTPDFRGKVFVASQRNLD 354
QY 341 SSTQNNPIVTVNSVKNSQFVPHLSITLTDENVSSGGDYIGTIQWTS--PPSDSG--GAN 396
Db 355 STTRAHEA---KVDTTAGRTPKLGSLEISTDSDDFOQOQTKFTPVGIGVDNEAE 407
QY 397 TNFWKIPDYGSSLAELASQAPVYPPGFNEVIVFYFNASIPGNQSG--SPNLVPCLLPQE 454
Db 408 FQWLSLPDYSGQFTHNMNLAPAVAPNFPGEQLLFRSOLP---SSGGRSNGVLDCLVPQE 464
QY 455 YITHFISQAPIQGEAALLHYVDPTNRNLGFEKLYPGGYLTCVPSNSSSTGQQQLPDGV 514
Db 465 WQHFYQESAPAQTOVALVRYVNPDTGKVLFEAKLHKLGFMTIANNGDS--PITVPPNGY 522
QY 515 VFASVWSRFFYQKPVGTAGPARGRL 540
Db 523 RFESWVNPFFYTLAPMGT--GNRRRI 547

RESULT 13
AAU91272
ID AAU91272 standard; protein; 548 AA.
XX AC AAU91272;
XX DT 18-JUN-2002 (first entry)
XX DE Norwalk virus associated polynucleotide #1.
XX KW Norwalk virus; monoclonal antibody; geno group I; geno group II;
XX KW immunological detection; food; viral infection.
XX OS Norwalk virus.
XX PN JP2002020399-A.
XX PD 23-JAN-2002.
XX PF 10-JUL-2000; 2000JP-00208151.
XX PR 10-JUL-2000; 2000JP-00208151.
XX PA (OSAP) OSAKA PREFECTURE.
XX PA (NISE-) NIPPON SEIBUTSU KAGAKU CENT KK.
XX PA (IATR) IATRON LAB INC.
XX DR WPI; 2002-287412/33.
XX PT A monoclonal antibody useful in the immunological detection and diagnosis
XX PT of Norwalk virus infection.
XX PS Disclosure; Page 12-13; 24pp; Japanese.
XX CC The invention describes a monoclonal antibody recognising Norwalk virus,
XX CC a capsid protein of Norwalk virus, or a common antigen epitope on the
XX CC capsid protein molecule of geno group I and geno group II. The antibody
XX CC is useful for immunological detection and quantitative analysis of
XX CC Norwalk virus in foods and the serum of infected patients. This sequence
XX CC represents a Norwalk virus associated protein described in the invention

XX SQ Sequence 548 AA;
Query Match 39.2%; Score 1134; DB 5; Length 548;
Best Local Similarity 44.5%; Pred No. 6e-92;
Matches 248; Conservative 87; Mismatches 180; Indels 42; Gaps 14;
QY 1 MMASKDAPTSADGATGAGQLVPEVNTADPIPIDPVAGSSTALATAGQVNLIDPMIINP 60
Db 1 MKMSSRAAPSNDGAAG--LVPEINN--EAMALDPVAGAAIAAPLTGQOIIIDPMINNF 56
QY 61 VOAPGEEITSPNNTPGDVLFDLOLQPHLNPLSHLSOMYNGMVGVMNRVVRVLAGNAFTA 120
Db 57 VOAPGEEVTVSPRNSPGVBLNLELGPBEINPYLAHARMYNGYAGGFEVQVVLAGNAFTA 116
QY 121 GKVIICUPPGFQSRITSLAQATLPHVIAVRLTDPVEVPLEDVRLVLYH--NNDTQPTM 179
Db 117 GKIIIFAAIPPNPIDNLSAAQITMCPHIVDVRLQEPVNLMPDVRRNFFHYNQSDSRL 176
QY 180 RLLCMLYPLRTGGASGGTDSFVAVAGVLTCPGDPNLFVLPPTVEOKTRPFTVPIPL 239
Db 177 RLIAMLYPLRLAN--NSGDDVFTVSCRVLTRPSDFSNFLVPPTVESKTKFTLPIITI 234
QY 240 KYLSNSRIPNPIEGMSLSPDOTQNVQFQNGRCTIDGQPLGTTTPVSVSOLCKFRGRI--- 296
Db 235 SEMSNRPFPVPIESLHTSPTENIVVQCQNGRVTLDELMTGTTQLLPSQICAPRGVLTST 294
QY 297 -----SGORVLN-----LTLDGSPF--MAFAAPAPAGFPDLGSCDWHIEMSKIPN 340
Db 295 SRASDAQATATPRLFNYYHVLQDLNLTGTPDPAEDI FGLGTPDFRGKVFVASQRNPD 354
QY 341 SSTQNNPIVTVNSVKNSQFVPHLSITLTDENVSSGGDYIGTIQWTSPPSDSG--GAN 397
Db 355 STTRAHEA---KVDTTAGRTPKLGSLEISTE--SSDFDONOQTRFT--PVGIGVDNEADF 408
QY 398 NFWKIPDYGSSLAELASQAPVYPPGFNEVIVFYFNASIPGNQSG--SPNLVPCLLPQE 455
Db 409 QQWLSLPDYSGQFTHNMNLAPAVAPNFPGEQLLFRSOLP---SSGGRSNGVLDCLVPQEW 465
QY 456 ITHFISQAPIQGEAALLHYVDPTNRNLGFEKLYPGGYLTCVPSNSSSTGQQQLPDGV 515
Db 466 WQHFYQESAPAQTOVALVRYVNPDTGKVLFEAKLHKLGFMTIANNGDS--PITVPPNGY 523
QY 515 VFASVWSRFFYQKPVGT 532
Db 524 RFESWVNPFFYTLAPMGT 540
RESULT 14
AAB49709
ID AAB49709 standard; protein; 550 AA.
XX AC AAB49709;
XX DT 04-APR-2001 (first entry)
XX DE Small round structured virus protein SEQ ID 10.
XX KW Small round structured virus; SRSV; food poisoning.
XX OS Small round structured virus.
XX PN WO2000079280-A1.
XX PD 28-DEC-2000.
XX PF 22-JUN-2000; 2000WO-JP004095.
XX PR 22-JUN-1999; 99JP-00175928.
XX PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX PA (DENK-) DENKA SEIKEN KK.

Wed Jun 2 09:13:28 2004

```

PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI; 2001-080848/09.
DR N-PSDB; AAF29150.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
XX Claim 1; Page 62-64; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 550 AA;
XX
Query Match 38.9%; Score 1124.5; DB 4; Length 550;
Best Local Similarity 44.2%; Pred. No. 4.3e-91;
Matches 250; Conservative 80; Mismatches 192; Indels 43; Gaps 14;
QY 1 MMASKDATPSADGATGAGQLVPEVNTADPITDPVAGSSTALATAGQVNLIDPWIINF 60
Db 1 MMASNDAPSDND--GAANLVEAND-EVMALEPVVGASIAAPVVGQNIIDPWIENF 56
QY 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSOMYNGWGNRVVVLGNAFTA 120
Db 57 VQAPQGEFTVSPRNSPCEMLNLELGPENLPHLSLRMYNGYAGQMQVVLGNAFTA 116
QY 121 GKVIICVPPGFSQRTLSIAQATLFPHVIAVTRTLDPEVEPLEDVRNVLYH--NNDTPTM 179
Db 117 GKIIFAAVPPHFPVENISAAQITWCHRVIVDVQLPVLPLDIRRFFHYNQENTPRM 176
QY 180 RLICMLYPLRTGASGSDTSFVAGRVLCPCGDPDNFLPVPPTVEQTRPFTVNIPL 239
Db 177 RLVAMLYPLR---ANSGEDVFTVSCRVLTRPAPDFEFTFLVPTVESKTKPFLPILTL 233
QY 240 KYLSNSRIPNPIEGMSLSPQTONVQFQNGRCTIDGQPLGTTTPVSVSQLCKFRGRI---- 295
Db 234 GELNSRFPAAIDMLYTDPNESIVVQFQNGRCTIDGLTQTTQVLTQICAFRGTLISQT 293
QY 296 -----TSGORVLN-----LTELDGSPP-MAFAAPAPAGFPDGLSCDWHIEMSKIPNS 341
Db 294 ARAADSTDSQARAHPLHVQVKNLDTQYDPTDDIEAVLCAIDFKGTVFGVASQRDVSG 353
QY 342 STONNPIVTVNS-----VKPNSQOFVPHLSITLDENVSSGGDYIGT---IQWTSPPSDSG 394
Db 354 QEEQHVATRAHEAHIDTDPKYAPKLGTL-----IKSGSDDFNTNQPIRFT--PVGMD 407
QY 395 ANTWFWKIPDYGSSLAELASOLAPAVYPGFNEVIVYFMASIPGNQSGSNLVPCLLPQE 454
Db 408 NNWQWELPDSIGRLTNLNLAPAVSPFQGERILFFRSIVPSAGGYGS-GYIDCLLPQE 466
QY 455 YTHFISBQAIQGSAAALLHYVDPDTRNLGELFKYPCGYLTCPVNSSSTGPOQLDGV 514
Db 467 WQHFYQEAAPSQSAVALRVYVNPDTGRNIFEAKLHREGFLTV--NCGNNPIVVPNGY 524
QY 515 FVPASWVRFRYQKVPVGTAGPARGR 539
Db 525 PRFEAWGNQFYTLAPMG-SGQRRR 548
XX
XX RESULT 15
XX ID AAB49710 standard; protein; 541 AA.
XX AC AAB49710;
XX DT 04-APR-2001 (first entry)

```

```

XX Small round structured virus protein SEQ ID 11.
XX
XX Small round structured virus; SRSV; food poisoning.
XX
XX Small round structured virus.
XX
XX WO200079280-A1.
XX
XX 28-DEC-2000.
XX
XX 22-JUN-2000; 2000WO-JP004095.
XX
XX 22-JUN-1999; 99JP-00175928.
XX
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI; 2001-080848/09.
XX N-PSDB; AAF29151.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX strains for investigation of food poisoning outbreaks, contains
XX antibodies.
XX
XX Claim 1; Page 64-66; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 541 AA;
XX
Query Match 33.6%; Score 970; DB 4; Length 541;
Best Local Similarity 38.0%; Pred. No. 2.6e-77;
Matches 224; Conservative 91; Mismatches 176; Indels 98; Gaps 17;
QY 1 MMASKATPSADGATGAGQLVPEVNTADPITDPVAGSSTALATAGQVNLIDPWIINF 60
Db 1 MMASNDAPSDNDGAAG---LVPEINN-EVMPFVAGASLATPVVQGOIIDPWIENN 56
QY 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSOMYNGWGNRVVVLGNAFTA 120
Db 57 VQAPAGEFTVSPRNSPGEILLDLLELGPENLPHLSLRMYNGYAGQMVEVQVILGNAFTA 116
QY 121 GKVIICVPPGFSQRTLSIAQATLFPHVIAVTRTLDPEVEPLEDVRNVLYH--NNDTPTM 179
Db 117 GKIIFAAVPPGFPVENISPSQITWCPHVIDVROLEPFLPEMPDIWNFFHYNQNDPKL 176
QY 180 RLICMLYPLRTGASGSDTSFVAGRVLCPCGDPDNFLVPPPTVEQTRPFTVNIPL 239
Db 177 RLVAMLYPLRAN--NSGDDVFTVSCRVLTKPSPDFEFTFLVPTVESKTKQFALPILKI 234
QY 240 KYLSNSRIPNPIEGMSLSPDQTONVQFQNGRCTIDGQPLGTTTPVSVSQLCKFRGR---- 294
Db 235 SEMTNSRFPVVDVVMYATARNENQVQPQNGRVTLGDELGLTTLPLAVNICFKGEVIAKN 294
QY 295 --ITSQORVLNLTELDGSPP-MAFAAPAPAGFPDGLSCDWHIEMSKIPNSSTONNPIVTN 351
Db 295 GSVRSYRMDMEITNTDGTPTDPTDPTGPGISGDFGILFGVASORKN---EQNP---- 347
QY 352 SVKPNQQFVPHLSITLDENVSSGGDYI-----GTIQWTSPP----- 388
Db 348 -----ATFAEALINTGDLCPQISSEIVLTSFNILRCTNPQPLPQSGLR 394
QY 389 -----PSDSGGAN-----TNFWKIPDYGSSLAELASOLAPAVYPGFNEVIVYF 431

```

Db 395 GTILIRSDNGHCHDMVGTSPPTTWPQQWRCRSGSNCCSSGHRYPV--PVMNRVTWIV 452
Qy 432 MASIFGPNGSGSPNLVPCLLPOEYITHFISEQAPIQGEAALLHYVDPDTNRNLGEFKLYP 491
Db 453 LSHKSGFSTSTRK-----LPQ-----LNLRW-----LIRFINPDIGRVLPEARLHK 494
Qy 492 GGYLTCTVPNSSSTGQQQLPLDGVFPASWVSRYOLKPVGTAGPARGRL 540
Db 495 QGFITVA--HTGDNFIVMPNGYFRFEAWVNOFYSLAPVGT-GKGRRRV 540

Search completed: June 1, 2004, 13:45:50
Job time : 48.7329 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 12.9611 Seconds
(without alignments)
2166.837 Million cell updates/sec

Title: US-09-926-799-4
Perfect score: 2891
Sequence: 1 MMASKDTPSADGATGAGQ.....YOLKPVGTAGPARGLGVR 544

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA*
1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCUS-COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2100	72.6	530	US-08-486-049-3	Sequence 3, Appli
2	268	9.3	668	US-09-617-594A-4	Sequence 4, Appli
3	266.5	9.2	626	US-09-590-020-7	Sequence 7, Appli
4	263.5	9.1	669	US-09-617-594A-2	Sequence 2, Appli
5	261.5	9.0	623	US-09-590-020-2	Sequence 2, Appli
6	261.5	9.0	623	US-09-590-020-4	Sequence 4, Appli
7	250.5	8.7	622	US-09-590-020-6	Sequence 6, Appli
8	143.5	5.0	2972	US-09-579-181-2	Sequence 2, Appli
9	143.5	5.0	3118	US-09-579-181-1	Sequence 1, Appli
10	138	4.8	2206	US-07-852-260-2	Sequence 2, Appli
11	138	4.8	2206	US-08-461-503-2	Sequence 2, Appli
12	138	4.8	2206	US-08-465-250-2	Sequence 2, Appli
13	121	4.2	2183	US-08-746-111-5	Sequence 5, Appli
14	119.5	4.1	1234	US-08-317-310A-15	Sequence 15, Appli
15	119.5	4.1	1234	PCT-US95-13041-15	Sequence 15, Appli
16	117.5	4.1	1155	US-08-094-948A-29	Sequence 29, Appli
17	117.5	4.1	1155	PCT-US96-09319-29	Sequence 29, Appli
18	113	3.9	1068	US-08-396-479B-12	Sequence 12, Appli
19	113	3.9	1068	US-08-818-823-12	Sequence 12, Appli
20	111.5	3.9	433	US-09-046-158A-2	Sequence 2, Appli
21	111.5	3.9	897	US-07-960-389-2	Sequence 2, Appli
22	111.5	3.9	2032	US-09-071-035-458	Sequence 458, App
23	111.5	3.9	2032	US-09-071-035-462	Sequence 462, App
24	111.5	3.9	2032	US-09-071-035-466	Sequence 466, App
25	111.5	3.9	2054	US-09-134-000C-6612	Sequence 6612, App
26	111	3.8	1075	PCT-US94-07297-41	Sequence 41, Appli
27	110	3.8	528	US-09-086-663A-82	Sequence 82, Appli

28	110	3.8	596	4	US-09-086-663A-2	Sequence 2, Appli
29	110	3.8	596	4	US-09-086-663A-80	Sequence 80, Appli
30	109	3.8	615	4	US-09-252-991A-26695	Sequence 26695, A
31	109	3.8	3892	4	US-09-328-352-5503	Sequence 5503, Ap
32	108.5	3.8	1911	4	US-09-854-856-64	Sequence 64, Appli
33	108.5	3.8	1939	4	US-09-854-856-48	Sequence 48, Appli
34	108.5	3.8	1971	4	US-09-854-856-32	Sequence 32, Appli
35	108.5	3.8	1999	4	US-09-854-856-16	Sequence 16, Appli
36	108.5	3.8	2004	4	US-09-854-856-58	Sequence 58, Appli
37	108.5	3.8	2032	4	US-09-854-856-42	Sequence 42, Appli
38	108.5	3.8	2048	4	US-09-854-856-62	Sequence 62, Appli
39	108.5	3.8	2064	4	US-09-854-856-26	Sequence 26, Appli
40	108.5	3.8	2076	4	US-09-854-856-46	Sequence 46, Appli
41	108.5	3.8	2092	4	US-09-854-856-10	Sequence 10, Appli
42	108.5	3.8	2108	4	US-09-854-856-30	Sequence 30, Appli
43	108.5	3.8	2136	4	US-09-854-856-14	Sequence 14, Appli
44	108.5	3.8	2141	4	US-09-854-856-56	Sequence 56, Appli
45	108.5	3.8	2157	4	US-09-854-856-52	Sequence 52, Appli

ALIGNMENTS

RESULT 1
US-08-486-049-3
; Sequence 3, Application US/08486049
; Patent No. 6572862
; GENERAL INFORMATION:
; APPLICANT: EG&G/Mary K
; APPLICANT: Jiang, Xi
; APPLICANT: Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Ave., N.W.
; CITY: Washington, D.C.
; STATE:
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,049
; FILING DATE: June 7, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311.023
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-486-049-3

Query Match
Best Local Similarity 72.6%
Score 2100; DB 4; Length 530;

Matches 396; Conservative 54; Mismatches 79; Indels 16; Gaps 8;

QY 1 MMASKDTPSADGATGAGQLVPEVNTADPDPIDVAGSSTALATAGQVNLDPWIINNF 60

DB 1 MMASKDTPSADGATGAGQLVPEVNTADPDPIDVAGSSTALATAGQVNLDPWIINNF 60

QY 61 VOAPQGEFTISNNTPGDLVLDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAGFTA 120
DB 61 VOAPQGEFTISNNTPGDLVLDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAGFTA 120
QY 121 GKVIICVPPGQSRSLTAQATLPPHVIADVRLDPEVEPLEDVRNVLHNND-TQPTM 179
DB 121 GKIIIVSCIPPGGSHNLITIAQATLPPHVIADVRLDPEVEPLEDVRNVLHNNDNQTM 180
QY 180 RLICMLYTEPLRTGGASGGTDSFVAVGRVLTCGPDNPLFLVPPTEVEKTPFTVBNPL 239
DB 181 RLVCMLYTEPLRTGGGTG--DSFVAVGRVMTCPSPDNFLFLVPPTEVEKTPFTLBNPL 238
QY 240 KYLSNSRIPIEGHSLSDPOTQVQFQNGRCTIDGQPLGTTPVSVSOLCKFRGRITSGQ 299
DB 239 SLSNSRAFLPISSMIGISPDNVQSVQFQNGRCTIDGRLVGTTPVSLHVAKIRG--TSNG 296
QY 300 RVNLNLTGDSPPMAFAAPAGFPDGLGSCDWHIEMSKIPNSSTQNNPIVTVSVKPNQQ 359
DB 297 TVNLNLTGDSPPHPEFAGPAGFPDGLGSCDWHIEMSKIPNSSTQNNPIVTVSVKPNQQ 351
QY 360 FVPHLSITLDENVSSGGDIYGTIQWTSPPSDSGGANTFWKIPDYGSSLAESOLAPAV 419
DB 352 FVPHLSI--QANGIGSGNYGVLSWISPPSHPSGQVDMKIPNGSSITEATHLAPSV 409
QY 420 YPGNEVIVFVMAISIPGPNQSGPNLVPCLLPQEIYTHFISEQAPIQGEAALLHYVDPD 479
DB 410 YPGFGEVLVFFNKMWGP--GAYNL-PCLLPQEIYTHLASEQAAPTVEAALLHYVDPD 465
QY 480 TNRNLGFKYPPGGLYTCVNSNSTGPOQLPDGVFVAFASWUSRFYOLKPVGTAGPARGR 539
DB 466 TGRNLGFKAYPDGFLTCVNGASGPOQLPFGVVFVSWVSRYQLKPVGTAGPARGR 525
QY 540 LGVR 544
DB 526 LGLR 530

RESULT 2
US-09-617-594A-4
; Sequence 4, Application US/09617594A
; Patent No. 6541458
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.1
; CURRENT APPLICATION NUMBER: US/09/617,594A
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-617-594A-4

Query Match 9.3%; Score 268; DB 4; Length 668;
Best Local Similarity 27.6%; Pred. No. 3.8e-16;
Matches 103; Conservative 54; Mismatches 140; Indels 76; Gaps 18;

QY 71 SPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAGFTAAGVTCVPP 130
DB 179 STSETQKILFKQSLGSLGPNLPLNLYTHLAKLVAVWMSGIEVRSISGSGVGGKLAIVVPP 238
QY 131 GFQS-RTLSIAQATLPPHVIADVRLDPEVEPLEDVRNVLHNNDTQPTMRLCMLYTP 189
DB 239 GIDPVQSTMLQ---YPHVLFDAQVEPFIPTDRLNLSLYHLMSDITTSILVINYDL 295

QY 190 RTGGASGGTDSFVAVGRVLTCGPDNPLFLVPPTEVEKTPFTVBNPIPLKYLNSRIPN 249
DB 296 INPYANDNSGCIIV-TVETKPGDPDFKHLKPPG-----SMLTHGSLPS 339
QY 250 PIEGMSLS-----PDQTONV---OFQ-NGRCTIDGQPLG-TTPVSVSOLCKFRGRI 295
DB 340 DLIPKSSSLWIGNRHSIDITDFVIKPFVQANRHFDFNOETAGWSTP-----RFR-PI 391
QY 296 TSGQRVLNLTGDSGPPMAFAAP-----APAGFPDGLGSCDWHIEMSKIPNSST-QNNPI 348
DB 392 T-----ITVSEKGGKLGIGVATDSIVPGIPDGMWD-----TTIPEKLIAPAGYA 436
QY 349 VTNVSKPNQQFVPHLSITLDENVSSGGDIY-GTIQ--WTSPPSDSGGANTFWKIPDY 405
DB 437 ITMGNNIDTTAADYDGASIIKNNTFKGYICGALQRAW-----GDKKISNTAF-----I 487
QY 406 GSSLAESOLAPA 418
DB 488 TTAIREGNSIKPS 500

RESULT 3
US-09-590-020-7
; Sequence 7, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vlinis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-7

Query Match 9.2%; Score 266.5; DB 4; Length 626;
Best Local Similarity 33.8%; Pred. No. 4.7e-16;
Matches 74; Conservative 32; Mismatches 88; Indels 25; Gaps 7;

QY 13 DGATGAGQLVPEVNT-ADPPIPDVAGSSTALATAGQVNLIDPWIINNFOAQOGEFTI- 70
DB 82 DGSITA---PEQGTWVGVIAEFSAQWSTAAADVATGKSVSEW-----EAFSFFH 128
QY 71 -----SPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAGFTAAGVTCVPP 125
DB 129 TSVNWSSTSETQKILFKQSLGSLGPNLPLNLYTHLAKLVAVWMSGIEVRSISGSGVGGKLA 188
QY 126 CCVPPGQFQS-RTLSIAQATLPPHVIADVRLDPEVEPLEDVRNVLHNNDTQPTMRLCML 184
DB 189 IVVPPGVDFVQSTMLQ---YPHVLFDAQVEPFIPTDRLNLSLYHLMSDITTSILVIM 245
QY 185 LYTPLRTGGASGGTDSFVAVGRVLTCGPDNPLFLVPP 223
DB 246 VYNDLINPYANDTNSGGCIV-TVETKPGDPDFKHLKPP 283

RESULT 4
US-09-617-594A-2
; Sequence 2, Application US/09617594A
; Patent No. 6541458
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT

```
; FILE REFERENCE: 454313-3151.1
; CURRENT APPLICATION NUMBER: US/09/617,594A
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-617-594A-2

Query Match          9.1%; Score 263.5; DB 4; Length 669;
Best Local Similarity 32.7%; Pred. No. 1.e-15;
Matches 72; Conservative 33; Mismatches 92; Indels 23; Gaps 7;

QY 12 ADGATGAGOLPEVNT-ADPIPIPVAGSSTALATAGQVNLIDPWIINNFOVQOGCEFTI 70
DB 125 ADGDSI--TTEQGTGVGVAESAQAATAAATGKSDSEW-----ESFESF 173

QY 71 -----SPNTPGDVFLDLQGLPHLPFLSHLSQMYNGVGNMRVVRVVLGNAGNAFTAGKVI 124
DB 174 HTSVNMTSETQKILFKQSLGPLNPLYLHLSKLYVAMSGSDVRFSGSGVFGKLA 233

QY 125 ICCVPGQGS-RTLSIAQATLPHVIADVRTLDPEVPLEDRVNLVHNNDTQPTMLLC 183
DB 234 AIVVPGVDVQVOSTSMLQ---YPHVLFDAQVPEVIFSDLRSTLYHLMSDITDTTSLVI 290

QY 184 MLYPLRTGGAGGTDTSFVAGRVLTCPGDPDFNLFVLP 223
DB 291 MYNDLNPYANDSNSSGCIV-TVETKPGDPFKFHLKPP 329

RESULT 5
US-09-590-020-2
; Sequence 2, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-2

Query Match          9.0%; Score 261.5; DB 4; Length 623;
Best Local Similarity 38.3%; Pred. No. 1.4e-15;
Matches 59; Conservative 27; Mismatches 63; Indels 5; Gaps 3;

QY 71 SPNTPGDVFLDLQGLPHLPFLSHLSQMYNGVGNMRVVRVVLGNAGNAFTAGKVIICVPP 130
DB 134 STSETQKILFKQSLGPLNPLYLHLSKLYVAMSGSVEVRSISGSGVFGKLAIVVPP 193

QY 131 GQGS-RTLSIAQATLPHVIADVRTLDPEVPLEDRVNLVHNNDTQPTMLLCMLYTPL 189
DB 194 GIEPVQSTSMQ---YPHVLFDAQVPEVIFAIPDLRNLNLYHLMSDITDTTSLVIMVYNDL 250

RESULT 6
US-09-590-020-4
; Sequence 4, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-4

Query Match          9.0%; Score 261.5; DB 4; Length 623;
Best Local Similarity 38.3%; Pred. No. 1.4e-15;
Matches 59; Conservative 27; Mismatches 63; Indels 5; Gaps 3;

QY 71 SPNTPGDVFLDLQGLPHLPFLSHLSQMYNGVGNMRVVRVVLGNAGNAFTAGKVIICVPP 130
DB 134 STSETQKILFKQSLGPLNPLYLHLSKLYVAMSGSVEVRSISGSGVFGKLAIVVPP 193

QY 131 GQGS-RTLSIAQATLPHVIADVRTLDPEVPLEDRVNLVHNNDTQPTMLLCMLYTPL 189
DB 194 GIEPVQSTSMQ---YPHVLFDAQVPEVIFAIPDLRNLNLYHLMSDITDTTSLVIMVYNDL 250

RESULT 7
US-09-590-020-6
; Sequence 6, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-6

Query Match          8.7%; Score 250.5; DB 4; Length 622;
Best Local Similarity 37.7%; Pred. No. 1.5e-14;
Matches 58; Conservative 27; Mismatches 64; Indels 5; Gaps 3;

QY 71 SPNTPGDVFLDLQGLPHLPFLSHLSQMYNGVGNMRVVRVVLGNAGNAFTAGKVIICVPP 130
```

Db 134 STSETQKILFKQSLGGLNPLNLSLAKLYVALAGSVEVRFSGSGVFGKLAIVVPP 193
Qy 131 GFQS-RTLSIAQATLPPHVIADVRTLPVPELDEDVRNVLHNNDTQPTWRLLCMLYTPL 189
Db 194 GIEPVQSTSMQLQ---YPHVLFDAQVEVIEFAIPDLRSNLYHLMSDITDTTSLVIMVYNDL 250
Qy 190 RTFGASGGTDSFVAVAGRVLTCPGDFNLFVPP 223
Db 251 INPYANDTSSGCIIV-TVETKPGDFKFKHLLKPP 283
RESULT 8
US-09-579-181-2
; Sequence 2, Application US/09579181
; Patent No. 6365372
; GENERAL INFORMATION:
; APPLICANT: Chivria, John
; APPLICANT: Yaciuk, Peter
; TITLE OF INVENTION: SNF2 Related CBP Activator Protein (SRCAP)
; FILE REFERENCE: 16153-4247
; CURRENT APPLICATION NUMBER: US/09/579,181
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/136,620
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2972
; TYPE: PRT
; ORGANISM: Human
US-09-579-181-2
Query Match 5.0%; Score 143.5; DB 4; Length 2972;
Best Local Similarity 22.1%; Pred. No. 0.0019;
Matches 124; Conservative 51; Mismatches 226; Indels 159; Gaps 22;
Qy 19 GQLVPEVNT-----ADPIDPVAGSS-TALATAGQVNLDPWLNINNFVQA-POGETTISP 72
Db 809 GRTVVVNNPRAPLGPVPRPPGPELSAQTPTGVPQVLP--ASLMVSASAPGPPILPA 866
Qy 73 NNTGDLVFDLQGLHNLNPLSHLSQMYNGWGNMR-----VRVVL 113
Db 867 SRPPGPPVLL-----PPQPNSGSLPQVLPSPGLVSGTSRPPPTLSLKPTTPAPVRLSP 921
Qy 114 AGNAFTAGKVIICCVPPGF-----QSRTLSIAQATLPPHVIADVRTLPDVEVPLED 164
Db 922 APPGSSLLKPLTVPPGCTTPPAAATTTTATATTTAVP-----APTAPQRLILSP 976
Qy 165 VRNVLYHNNDTQPTWRLLCMLYPLRTGASGGTDSFVAVAGRVLTCPGDFNLFVLP----- 220
Db 977 DMQARLPSEVVSIGQLASLAQRVANAGGSKPL-TFQIQGNKLTLTGAQVRLAVGQPR 1035
Qy 221 ---VPPT-----VEQTRPFTVPTNPL-----KYLNSGRIP 248
Db 1036 PLQMPPTMNNVTGVVKIVVRQAPRDLTPVPPLAPAPRPPSSGLPAVLNPRPTLTPGRLP 1095
Qy 249 NPIEGMSLSPDQTNQVQNGRCTIDGQPLGTTTPVSVSQLCKFRGRTTSGORVNLNLTELD 308
Db 1096 TPTLGTARAP-----MP---TPTLVRLPK-----LV 1119
Qy 309 GSPFMAFAAPAPAGPDLGSCDWHIEMS-----KIPNSSTQNNPIVTVNSVKPNSQQ 359
Db 1120 HSPSPVVSASAPGAAPLTITSSPLHVPSSLPGPASSPMPIPNSPLASVSVSVPLSSS 1179
Qy 360 F-----VPHLSITLDENVSSGGYIGTIQWTSF-----PSDSCGANTNEWK 401
Db 1180 LPISVPTTLPAAPASAPLTIPISAPLTVSASGALLT-SVTPLAPVVPAAAPGPPSLQ--- 1235
Qy 402 IPDYGSSLABASQALAVYP-----PGFNEVIVVFMAIIPGNOSGSPNLVPCILLPOE 454
Db 1236 -PSGASPSASALTGLATAPSLSSSQTPGHPHLLAPTSSHVPLGLNSTVAPACSPVLVPAS 1294

Qy 455 YI-THFISEQAPIQGEAALL 473
Db 1295 ALASPPSPAPNAPAPQAASLL 1314
RESULT 9
US-09-579-181-1
; Sequence 1, Application US/09579181
; Patent No. 6365372
; GENERAL INFORMATION:
; APPLICANT: Chivria, John
; APPLICANT: Yaciuk, Peter
; TITLE OF INVENTION: SNF2 Related CBP Activator Protein (SRCAP)
; FILE REFERENCE: 16153-4247
; CURRENT APPLICATION NUMBER: US/09/579,181
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/136,620
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3118
; TYPE: PRT
; ORGANISM: Human
US-09-579-181-1
Query Match 5.0%; Score 143.5; DB 4; Length 3118;
Best Local Similarity 22.1%; Pred. No. 0.0021;
Matches 124; Conservative 51; Mismatches 226; Indels 159; Gaps 22;
Qy 19 GQLVPEVNT-----ADPIDPVAGSS-TALATAGQVNLDPWLNINNFVQA-POGETTISP 72
Db 955 GRTVVVNNPRAPLGPVPRPPGPELSAQTPTGVPQVLP--ASLMVSASAPGPPILPA 1012
Qy 73 NNTGDLVFDLQGLHNLNPLSHLSQMYNGWGNMR-----VRVVL 113
Db 1013 SRPPGPPVLL-----PPQPNSGSLPQVLPSPGLVSGTSRPPPTLSLKPTTPAPVRLSP 1067
Qy 114 AGNAFTAGKVIICCVPPGF-----QSRTLSIAQATLPPHVIADVRTLPDVEVPLED 164
Db 1068 APPGSSLLKPLTVPPGCTTPPAAATTTTATATTTAVP-----APTAPQRLILSP 1122
Qy 165 VRNVLYHNNDTQPTWRLLCMLYPLRTGASGGTDSFVAVAGRVLTCPGDFNLFVLP----- 220
Db 1123 DMQARLPSEVVSIGQLASLAQRVANAGGSKPL-TFQIQGNKLTLTGAQVRLAVGQPR 1181
Qy 221 ---VPPT-----VEQTRPFTVPTNPL-----KYLNSGRIP 248
Db 1182 PLQMPPTMNNVTGVVKIVVRQAPRDLTPVPPLAPAPRPPSSGLPAVLNPRPTLTPGRLP 1241
Qy 249 NPIEGMSLSPDQTNQVQNGRCTIDGQPLGTTTPVSVSQLCKFRGRTTSGORVNLNLTELD 308
Db 1242 TPTLGTARAP-----MP---TPTLVRLPK-----LV 1265
Qy 309 GSPFMAFAAPAPAGPDLGSCDWHIEMS-----KIPNSSTQNNPIVTVNSVKPNSQQ 359
Db 1266 HSPSPVVSASAPGAAPLTITSSPLHVPSSLPGPASSPMPIPNSPLASVSVSVPLSSS 1325
Qy 360 F-----VPHLSITLDENVSSGGYIGTIQWTSF-----PSDSCGANTNEWK 401
Db 1326 LPISVPTTLPAAPASAPLTIPISAPLTVSASGALLT-SVTPLAPVVPAAAPGPPSLQ--- 1381
Qy 402 IPDYGSSLABASQALAVYP-----PGFNEVIVVFMAIIPGNOSGSPNLVPCILLPOE 454
Db 1382 -PSGASPSASALTGLATAPSLSSSQTPGHPHLLAPTSSHVPLGLNSTVAPACSPVLVPAS 1440
Qy 455 YI-THFISEQAPIQGEAALL 473
Db 1441 ALASPPSPAPNAPAPQAASLL 1460
RESULT 10
US-07-852-260-2

; Sequence 2, Application US/07852260
; Patent No. 5525715
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/852,260
; FILING DATE: 19920619
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-B-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-852-260-2

Query Match 4.8%; Score 138; DB 1; Length 2206;
Best Local Similarity 24.2%; Pred. No. 0.0038;
Matches 61; Conservative 31; Mismatches 110; Indels 50; Gaps 10;

QY 22 VPEVNTADPIPIPVAGSSTALATAGQVNLDPWLNINNVQAPQG-----EFTISPNTTP 76
Db 365 IPEFVTPPIDIPGVKNMELAE-----IDTMPLNLESTKRTNMDMYRVTLSDSADL 418

QY 77 GDLVFDLQLGPHLNPFLSH-----LSQMYNGWGNMVRVVLNAGNAFTAGKVIICCVPPG 131
Db 419 SQPILCLSLSPAFDPRLSHTMLGEVLNYYTHWAGSLKFTFLFCGSMMATGKILVAYAPG 478

QY 132 FQSRTLSIAQATLPHVIAVTRTLDPEVPLEDVNRVLYHNNDTPTMR---LLCWLX-- 186
Db 479 AQPPT-SRKEAMLGTHVINDLGQSCTMVVWVWISNVTYRQT-TQDSFTEGGYISMFYQT 536

QY 187 ---TPLRTGGASGTDTSFVWAGRVLTCPGDPNFLEFLVPTVEQKTRPTVTNIPKYL 243
Db 537 RIWVPLST-----PKSMSMLGFVSAC--NDFSRLRLDRTT-----HIS 572

QY 244 NSRIENPIEGMS 255
Db 573 QSALPQGIEDLT 584

RESULT 11
US-08-461-503-2
; Sequence 2, Application US/08461503
; Patent No. 5834302
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.

; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES
; TITLE OF INVENTION: FROM CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,503
; FILING DATE: 5-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-D-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-461-503-2

Query Match 4.8%; Score 138; DB 2; Length 2206;
Best Local Similarity 24.2%; Pred. No. 0.0038;
Matches 61; Conservative 31; Mismatches 110; Indels 50; Gaps 10;

QY 22 VPEVNTADPIPIPVAGSSTALATAGQVNLDPWLNINNVQAPQG-----EFTISPNTTP 76
Db 365 IPEFVTPPIDIPGVKNMELAE-----IDTMPLNLESTKRTNMDMYRVTLSDSADL 418

QY 77 GDLVFDLQLGPHLNPFLSH-----LSQMYNGWGNMVRVVLNAGNAFTAGKVIICCVPPG 131
Db 419 SQPILCLSLSPAFDPRLSHTMLGEVLNYYTHWAGSLKFTFLFCGSMMATGKILVAYAPG 478

QY 132 FQSRTLSIAQATLPHVIAVTRTLDPEVPLEDVNRVLYHNNDTPTMR---LLCWLX-- 186
Db 479 AQPPT-SRKEAMLGTHVINDLGQSCTMVVWVWISNVTYRQT-TQDSFTEGGYISMFYQT 536

QY 187 ---TPLRTGGASGTDTSFVWAGRVLTCPGDPNFLEFLVPTVEQKTRPTVTNIPKYL 243
Db 537 RIWVPLST-----PKSMSMLGFVSAC--NDFSRLRLDRTT-----HIS 572

QY 244 NSRIENPIEGMS 255
Db 573 QSALPQGIEDLT 584

RESULT 12
US-08-465-250-2
; Sequence 2, Application US/08465250
; Patent No. 6136570
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:


```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/317,310A
; FILING DATE: 03-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: JDP-022
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1234 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-317-310A-15

```

```

Query Match          4.1%; Score 119.5; DB 2; Length 1234;
Best Local Similarity 20.5%; Pred. No. 0.079;
Matches 98; Conservative 68; Mismatches 140; Indels 171; Gaps 29;

QY 156 DVEVPLEDRVNLVHNNDTQTMRLLCMLYPLRTGASGCTDSFVWAGRVLTCPGPDF 215
DB 273 NPISVPLR-----RHLLNPPPSQ---VGLTRRSRSTESITATSPASWVGK-----PGS--- 318
QY 216 NLFVLVPTVEOK---TRPFTVNIPL-----KYLNSRIPNPIEGMSLSPDQTO 262
DB 319 ---FRVRASDGEGTMSRPASVDGSPVSPSTNRTHAHRHSGSLRHPPL-----NHSR 368
QY 263 NVQFONGRCTIDQPLGTTTPVSQSL-----CKFRGRITSGORVNLNLTDLGSPF 312
DB 369 SIPMPSRCS---PSATSPVLSLSSSTSGHGSTDCFLPRSSA-----SVSGS--- 414
QY 313 MAFAPAPAGF---PDLGS---CDWHIEMSKIIPNSSTQNNPIVTNSVKPNSQQFVPHLSI 367
DB 415 -----PSDGGFISSDEYSGSPCDFRSSFRSVTPDLSLHTPPARGEEELSNYICMGKGAS 469
QY 368 TLDEN-----VSSGSD---YI-GTIQWTSF---PSD---SGGAN-----T 397
DB 470 TLAAPNGHYILSRGGNGHRYIPGANLGTSPALPGDEAAGADLNRFRKRTHSAGTSPTI 529
QY 398 NFWKIPDYGs---SLAASQAPAVYPPGFNEVIVYFMASIPGPNQSGSNLVPCLLPQBY 455
DB 530 SHQKTPSQSSVASIEEYTEMPPAAYPPG-----GGSGRLPGYRHSAP 572
QY 456 I-THFISQAPIQGEAALLHYVD-----PDTNRNL-----GEFKLYP----- 491
DB 573 VTHSYPE-----EGLEMHLLERRGGHRRPDTs-NLHTDDGYMPMSQVAPVPSNRKGN 625
QY 492 GGYLTCVPNNSST-----GQQLPLDGVFVFWASVWSRYQLKPVGTAGPARG 538
DB 626 GGYMPMSPKSVSAPOQIINPIRRHPQVRDPNG-----YMMWSPSGSCSPDIG 672

```

RESULT 15

```

PCT-US95-13041-15
; Sequence 15, Application PC/TUS9513041
; GENERAL INFORMATION:
; APPLICANT: WHITE, Morris F.
; APPLICANT: SUN, Xiao Jian
; APPLICANT: PIERCE, Jacalyn H.
; TITLE OF INVENTION: THE IRS FAMILY OF GENES
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts

```

```

; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13041
; FILING DATE: Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/317,310
; FILING DATE: 03-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: JDP-022PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1234 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; PCT-US95-13041-15

```

```

Query Match          4.1%; Score 119.5; DB 5; Length 1234;
Best Local Similarity 20.5%; Pred. No. 0.079;
Matches 98; Conservative 68; Mismatches 140; Indels 171; Gaps 29;

QY 156 DVEVPLEDRVNLVHNNDTQTMRLLCMLYPLRTGASGCTDSFVWAGRVLTCPGPDF 215
DB 273 NPISVPLR-----RHLLNPPPSQ---VGLTRRSRSTESITATSPASWVGK-----PGS--- 318
QY 216 NLFVLVPTVEOK---TRPFTVNIPL-----KYLNSRIPNPIEGMSLSPDQTO 262
DB 319 ---FRVRASDGEGTMSRPASVDGSPVSPSTNRTHAHRHSGSLRHPPL-----NHSR 368
QY 263 NVQFONGRCTIDQPLGTTTPVSQSL-----CKFRGRITSGORVNLNLTDLGSPF 312
DB 369 SIPMPSRCS---PSATSPVLSLSSSTSGHGSTDCFLPRSSA-----SVSGS--- 414
QY 313 MAFAPAPAGF---PDLGS---CDWHIEMSKIIPNSSTQNNPIVTNSVKPNSQQFVPHLSI 367
DB 415 -----PSDGGFISSDEYSGSPCDFRSSFRSVTPDLSLHTPPARGEEELSNYICMGKGAS 469
QY 368 TLDEN-----VSSGSD---YI-GTIQWTSF---PSD---SGGAN-----T 397
DB 470 TLAAPNGHYILSRGGNGHRYIPGANLGTSPALPGDEAAGADLNRFRKRTHSAGTSPTI 529
QY 398 NFWKIPDYGs---SLAASQAPAVYPPGFNEVIVYFMASIPGPNQSGSNLVPCLLPQBY 455
DB 530 SHQKTPSQSSVASIEEYTEMPPAAYPPG-----GGSGRLPGYRHSAP 572
QY 456 I-THFISQAPIQGEAALLHYVD-----PDTNRNL-----GEFKLYP----- 491
DB 573 VTHSYPE-----EGLEMHLLERRGGHRRPDTs-NLHTDDGYMPMSQVAPVPSNRKGN 625
QY 492 GGYLTCVPNNSST-----GQQLPLDGVFVFWASVWSRYQLKPVGTAGPARG 538
DB 626 GGYMPMSPKSVSAPOQIINPIRRHPQVRDPNG-----YMMWSPSGSCSPDIG 672

```

Search completed: June 1, 2004, 13:57:57
Job time : 13.9611 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 34.7758 Seconds
(without alignments)
4368.312 Million cell updates/sec

Title: US-09-926-799-4
Perfect score: 2891
Sequence: 1 MMASKATPSADGATGAGQ.....YOLKPVGTAGARGRLGVR 544

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2100	72.6	530	14	US-10-314-739-3
2	268	9.3	668	14	Sequence 3, Appli
3	264.5	9.1	547	12	Sequence 4, Appli
4	264.5	9.1	671	12	Sequence 4, Appli
5	263.5	9.1	669	14	Sequence 2, Appli
6	143.5	5.0	2971	14	Sequence 2, Appli
7	130	4.5	3930	12	Sequence 50, Appl
8	126.5	4.4	1185	12	Sequence 46817, A
9	125.5	4.3	2358	12	Sequence 46893, A
10	123.5	4.3	6310	12	Sequence 45763, A
11	119.5	4.1	1231	16	Sequence 67793, A
12	117.5	4.1	609	16	Sequence 3, Appli
13	117	4.0	1564	15	Sequence 43, Appl
14	114.5	4.0	913	15	Sequence 1254, Ap
15	114.5	4.0	1212	16	Sequence 2294, Ap
					Sequence 5, Appli

16	113.5	3.9	618	14	US-10-061-201-7
17	113.5	3.9	729	14	US-10-061-201-3
18	113.5	3.9	1147	15	US-10-327-481A-38
19	113	3.9	709	12	US-10-231-956A-287
20	113	3.9	910	16	US-10-389-566-2431
21	113	3.9	1794	10	US-09-985-738-299
22	113	3.9	1799	10	US-09-985-738-149
23	113	3.9	1821	10	US-09-965-738-82
24	113	3.9	2234	12	US-10-612-090-20
25	113	3.9	5877	14	US-10-142-515-11
26	113	3.9	5935	14	US-10-243-243A-8
27	113	3.9	11721	10	US-09-965-738-162
28	112	3.9	1023	12	US-10-147-493-200
29	112	3.9	1023	12	US-10-145-127-200
30	112	3.9	1023	12	US-10-160-503-200
31	112	3.9	1023	12	US-10-143-118-200
32	112	3.9	1023	12	US-10-144-993-200
33	112	3.9	1023	12	US-10-158-787-200
34	112	3.9	1023	12	US-10-140-024-200
35	112	3.9	1023	12	US-10-140-808-200
36	112	3.9	1023	12	US-10-152-405-200
37	112	3.9	1023	12	US-10-137-852A-200
38	112	3.9	1023	12	US-10-127-900A-200
39	112	3.9	1023	12	US-10-128-685A-200
40	112	3.9	1023	12	US-10-131-820A-200
41	112	3.9	1023	12	US-10-142-886-200
42	112	3.9	1023	12	US-10-146-728-200
43	112	3.9	1023	12	US-10-146-786-200
44	112	3.9	1023	12	US-10-147-499-200
45	112	3.9	1023	12	US-10-157-798-200

ALIGNMENTS

RESULT 1

US-10-314-739-3
; Sequence 3, Application US/10314739
; Publication No. US20030129588A1
; GENERAL INFORMATION:
; APPLICANT: ~~Estes, Mary K~~
; Jiang, Xi
; Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; Characterize No. US20030129588A1walk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; CITY: Washington, D.C.
; STATE: <Unknown>
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/314,739
; FILING DATE: 09-Dec-2002
; CLASSIFICATION: <Unknown>
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,049
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311,023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX: <Unknown>

Wed Jun 2 09:13:28 2004

```
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 530 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-314-739-3

Query Match          72.6%; Score 2100; DB 14; Length 530;
Best Local Similarity 72.7%; Pred. No. 7.2e-182;
Matches 396; Conservative 54; Mismatches 79; Indels 16; Gaps 8;

QY 1 MMWASKDAPSDAGATGAGQLVPEVNTADPIPIDPVAGSSTALATAGQVNLDPWIINNF 60
DB 1 MMWASKDAPSDVAGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPDIWIINNF 60
QY 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAFTA 120
QY 121 GKVIICVPPGFOSRSLTAQATLFEHVIADVRTLDPEVEPLEDVNRVLYHNND-TQPTM 179
DB 121 GKVIICVPPGFOSRSLTAQATLFEHVIADVRTLDPEVEPLEDVNRVLYHNND-TQPTM 180
QY 121 GKVIICVPPGFOSRSLTAQATLFEHVIADVRTLDPEVEPLEDVNRVLYHNND-TQPTM 179
DB 121 GKVIICVPPGFOSRSLTAQATLFEHVIADVRTLDPEVEPLEDVNRVLYHNND-TQPTM 180
QY 180 RLKLCMLYPLRTGASGGTDSFVAVAGRLTCGPDNFNLFVPPPTVEQKTRPFTVPNPL 239
DB 181 RLKLCMLYPLRTGAGTG--DSFVAVAGRLTCGPDNFNLFVPPPTVEQKTRPFTVPNPL 238
QY 240 KYLSNSRIPNIPGMSLSPPQTQNVQFONGRCTIDGQPLGTTTPVSVSQLCKFRGRTSQ 299
DB 239 SLSNSRAPLPISSMGISPDNVQSVQFONGRCTLDGRLVGTTPVSVSHVAKING--TSNG 296
QY 300 RVNLNLTDELGSPFMAAPAPAGPDLGSCDHIEMSKIPNSSTQNNPIVTVNSVKPNSQ 359
DB 297 TVNLNLTDELGTPFPFPGFAPIGFDPDLGGCDHINMTQGHSSQTYQVDVDTT---PDT-- 351
QY 360 FVPHLSITLDENVSSGGDYIGTIQWTSPPSDSGGANTNFWKIPDYGSSLAELASQAPAV 419
DB 352 FVPHLSI--QANGIGSGNTVGVLSWISPPSHSPSGQVLDLWKIPNGSSITEATHLAPSV 409
QY 420 YPPGFNEVIVFMASIRGPNQSGSNLPVCLLPQEIYTHISEQAPIQGEAALLHYVDPD 479
DB 410 YPPGFGEVIVFMASIRGPNQSGSNLPVCLLPQEIYTHISEQAPIQGEAALLHYVDPD 465
QY 480 TNRNLGEPKLYPGGYLICVPSNSSSTGQPLDGVFVAFSVKRSFYOLKPVGTAGPARGR 539
DB 466 TGRNLGEPKLYPGGYLICVPSNSSSTGQPLDGVFVAFSVKRSFYOLKPVGTAGPARGR 525
QY 540 LGVRR 544
DB 526 LGLRR 530

RESULT 2
US-10-209-507-4
; Sequence 4, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-4

Query Match          9.3%; Score 268; DB 14; Length 668;
Best Local Similarity 27.6%; Pred. No. 4.3e-15;
Matches 103; Conservative 54; Mismatches 140; Indels 76; Gaps 18;

QY 71 SPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAFTAGKVIICVPP 130
DB 179 STSETQKILFKQSLGPLENPNLTHLAKLYAVMSGIEVRSISGSGVFGGKLAIVPP 238
QY 131 GFOS-RTLSIAQATLFEHVIADVRTLDPEVEPLEDVNRVLYHNNDTQPTMRLLCMLYTP 189
DB 239 GIDPVQSTMLQ---YPHVLFDARQVEPVIFIPDLRNSLYHLMSDPTDTSVLMIYNDL 295
QY 190 RTGASGGTDSFVAVAGRLTCGPDNFNLFVPPPTVEQKTRPFTVPNIPKLYLSNRIPN 249
DB 296 INPYANDSNSGCIV-TVETKPGDPFKFHLKPPG-----SMLTHGSIPS 339
QY 250 PIEGMSLS-----PDQTQNV-----QFQ-NGRCTIDGQPLG-TTPSVSQLCKFRGRI 295
DB 340 DLIPKSSSLWIGNRHMSDIITDFVFKPFVQANRHDFNOETAGWSTP-----REF-PI 391
QY 296 TSGORVLNLTDELGSPFMAAP-----APAGFPDLGSCDHIEMSKIPNSST-QNNPI 348
DB 392 T-----ITVSEKSGSLGIGVATDSIVPGIPDCMPD-----TTIPEKLTAPAGDYA 436
QY 349 VTVNSVKPNSQFVPHLSITLDENVSSGGDYI-GTIQ--WTSPPSDSGGANTNFWKIPDY 405
DB 437 ITNGGNNDITTAADYDGASIIKNTNPKGMYICGALQRAW-----GDKKISNTAF-----I 487
QY 406 GSSLAELASQAPAV 418
DB 488 TTAIREGNSIKPS 500

RESULT 3
US-10-670-695-4
; Sequence 4, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-670-695-4

Query Match          9.1%; Score 264.5; DB 12; Length 547;
Best Local Similarity 33.3%; Pred. No. 6.6e-15;
Matches 73; Conservative 32; Mismatches 89; Indels 25; Gaps 7;

QY 13 DGATGAGQLVPEVNT-ADPIPIDPVAGSSTALATAGQVNLDPWIINNFVQAPQGEFTI- 70
DB 3 DGSITA-----PEQGTWVGVIAPESQAMSTAAADMATKGSVDSEW-----EAFPSFH 49
QY 71 -----SPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAFTAGKVI 125
```

Db 50 TSNVSTSTQKILFKQSLGELLNPYLEHLAKLYVAMSGSIEVRFSISGSGVFGKLA 109
Qy 126 CCVPPGFQS-RTLSIAQATLFPFHVIADVRTLDPEVEPLEDVRNVLVYHNNNDTQPTMRLLCM 184
Db 110 IVVPPGVDPVQSTSLQ---YPHVLFDAQVEPVIFCLPDLASTLYHLMSDITDTTSLVIM 166
Qy 185 LYTPLRTGASGSGTDFVAVGRVLCPCGPDNFLELVPP 223
Db 167 VYNDLINPYANDANSSGCIV-TVETKPGDPDFKFLKPP 204

RESULT 4

US-10-670-695-2
; Sequence 2, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 671
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-670-695-2

Query Match 9.1%; Score 264.5; DB 12; Length 671;
Best Local Similarity 33.3%; Pred. No. 9.1e-15;
Matches 73; Conservative 32; Mismatches 89; Indels 25; Gaps 7;
Qy 13 DGATGAGQLVPEVNT-ADPIPDPVAGSSTALATAGQVNLIDPWIINNFPVQAPQGEFTI- 70
Db 127 DGSITA----PQGTWVGVIAPSAQMSAADMATGKSDSEW-----EAFSPH 173
Qy 71 ----SPNTPGDVLFDLQGLPHLPFLSHLSQMYNGWGNMRVRVVLGNAFTAGKVII 125
Db 174 TSNVSTSTQKILFKQSLGELLNPYLEHLAKLYVAMSGSIEVRFSISGSGVFGKLA 233
Qy 126 CCVPPGFQS-RTLSIAQATLFPFHVIADVRTLDPEVEPLEDVRNVLVYHNNNDTQPTMRLLCM 184
Db 234 IVVPPGVDPVQSTSLQ---YPHVLFDAQVEPVIFCLPDLASTLYHLMSDITDTTSLVIM 290
Qy 185 LYTPLRTGASGSGTDFVAVGRVLCPCGPDNFLELVPP 223
Db 291 VYNDLINPYANDANSSGCIV-TVETKPGDPDFKFLKPP 328

RESULT 5

US-10-209-507-2
; Sequence 2, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11

; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-2

Query Match 9.1%; Score 263.5; DB 14; Length 669;
Best Local Similarity 32.7%; Pred. No. 1.1e-14;
Matches 72; Conservative 33; Mismatches 92; Indels 23; Gaps 7;
Qy 12 ADGATGAGQLVPEVNT-ADPIPDPVAGSSTALATAGQVNLIDPWIINNFPVQAPQGEFTI 70
Db 125 ADGSSSI--TTPEQGTLVGGVIAEPSAQMATADAATGKSDSEW-----SSFFSF 173
Qy 71 ----SPNTPGDVLFDLQGLPHLPFLSHLSQMYNGWGNMRVRVVLGNAFTAGKVI 124
Db 174 HTSVNMSTSTQKILFKQSLGELLNPYLEHLAKLYVAMSGSVDVRFISGSGVFGKLA 233
Qy 125 ICCVPPGFQS-RTLSIAQATLFPFHVIADVRTLDPEVEPLEDVRNVLVYHNNNDTQPTMRLLC 183
Db 234 AIIVPPGVDPVQSTSLQ---YPHVLFDAQVEPVIFSIPLRLSTLYHLMSDITDTTSLVI 290
Qy 184 MLYTPLRTGASGSGTDFVAVGRVLCPCGPDNFLELVPP 223
Db 291 VYNDLINPYANDANSSGCIV-TVETKPGDPDFKFLKPP 329

RESULT 6

US-10-146-473-50
; Sequence 50, Application US/10146473
; Publication No. US2003010888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 50
; LENGTH: 2971
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-473-50

Query Match 5.0%; Score 143.5; DB 14; Length 2971;
Best Local Similarity 22.1%; Pred. No. 0.009;
Matches 124; Conservative 51; Mismatches 226; Indels 159; Gaps 22;
Qy 19 GOLVPEVNT----ADPIPDPVAGSS-TALATAGQVNLIDPWIINNFPVQAPQGEFTISP 72
Db 809 GRTVVVNNRAPLGPVDPVRPPGPPELSAQTPGPGVQVLP--ASLMVSASAPGPPPIPA 866
Qy 73 NNTPGDVLFDLQGLPHLPFLSHLSQMYNGWGNMR-----VRVVL 113
Db 867 SRPPGPVLL-----PPLQPNSGSLPQVLPSPGLVLSGTSRPPPTPLSLKPTPPAPVRUSP 921
Qy 114 AGNAFTAGKVIICCPVPGF-----QSRTLSTIAQATLFPFHVIADVRTLDPEVEPLED 164
Db 922 APPGSSSLAKPLTVPPGYTFPPAAATTTTATATTAVP-----APTAPQQLIISP 976
Qy 165 VRNVLVYHNNNDTQPTMRLLCMLYTLPLRTGASGSGTDFVAVGRVLCPCGPDNFLELV 220

Db 977 DQARLPGEVVSGVIGLALSLAQRVANAGGSKPL-TFQIQGNKLTITLGAQVROLAVGQPR 1035
Qy 221 ---VPPT-----VEQTRFTVPNIPL-----KYLNSRIP 248
Db 1036 PLOMPTVNNVTGVVQVQNGRECTIDGOLGTTTPVVSQCLCKFRGRTTSGORVNLNLTELD 308
Qy 249 NPIEGMSLSDDQNTQVQNGRECTIDGOLGTTTPVVSQCLCKFRGRTTSGORVNLNLTELD 308
Db 1096 TPTLGTARAP-----MP-----TPTLVRPLK-----LV 1119
Qy 309 GSPFMAFAAPAPAGPDGSCDWHIEMS-----KIPNSSTONNPIVNSVKPNQQ 359
Db 1120 HSPFVSASACAPLTISSPLHVPSSLPGPASSPMPINSSPLASPVSTVSVPLSSS 1179
Qy 360 F-----VPHLSITLDENVSSGGDYICQIOWTSP-----PSDSGGANTNFWK 401
Db 1180 LPISVPTTLPAPASAPLTIPISAPLTVSASGALLT-SVTPPLAPVVPAAAGPPLQ--- 1235
Qy 402 IPDYGSSLAELASQAFAPVVP-----PGFNEVIVVFMAISIPGPNOSGPNLVPCLLPOE 454
Db 1236 -PSGAPSPASALTGLGATAPSLSSSQTPGHPHLLAPTSSHVPLGNSVAPACSPVLVPAS 1294
Qy 455 YI-THFISEQAPIQSGAALL 473
Db 1295 ALASPPPSAPNPAPAQASLL 1314

RESULT 7
US-10-282-122A-46817
; Sequence 46817, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haeelbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46817

LENGTH: 3930
TYPE: PRT
ORGANISM: Bacillus anthracis
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (19)..(19)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (38)..(38)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (3455)..(3455)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (3459)..(3459)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (3552)..(3552)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (3566)..(3566)
OTHER INFORMATION: X=any amino acid
US-10-282-122A-46817

Query Match 4.5%; Score 130; DB 12; Length 3930;
Best Local Similarity 21.1%; Pred. No. 0.23;
Matches 136; Conservative 74; Mismatches 217; Indels 218; Gaps 33;
Qy 4 ASKDATPSADGATGA-QQLVPEVNTA-----DPIDPVGAGSTALATAGQVNLID 53
Db 1506 ASVIATKTASSAFAAAGDTITTTTLTNSGNTTANTFVFDILPAELSFVDSVQINTI- 1564
Qy 54 PWIINNFVQAPQ-----PQLGFRPDTGVPLDSIPVGGTITISFOAIVGSI- 1607
Db 1565 -----PQLGFRPDTGVPLDSIPVGGTITISFOAIVGSI- 1607
Qy 97 SOMYNGWVGNMR--VRVVLGNAGFTAGKVIICCVPGFQSRRTLSIAQAT-----LPHVI 149
Db 1608 STTYSIIIVDTPQPPVTEIATSN-----PTLIQINEAITQATKSVDRLFSDV- 1653
Qy 150 ADVRTLDDEVEPLDEDVNLVYHN-NDQPTMRLCLMLYPLRTGG-----ASGTD 200
Db 1654 APGNSFLTYTLLNIGNTTATNIIFTDPIPNHTVFIEDSVRVGILLPGVNPANGIPG 1713
Qy 201 FVWAG-----RVLTCPGPDF-----NFLVLPVPTVEQTRFTV 234
Db 1714 DIIAGDFINITFRVQVVSIPNPIFTIGPGPNSPVWNGASINYQFM-----TG 1761
Qy 235 FNIPLKLSNRPINPIEGMSLSDDQNTQVQNGRECTIDGOLGTTTPVVSQCLCKFR- 292
Db 1762 FNLPL-ASRSTTSNEV-----STQINSGEIAL-VKSDVKFTVTIGDITLSYGIS 1807
Qy 293 -----GRITSGORVNLNLTGDSPEMAFAAPAGPDGSCDWHIEMSKI PNSTQN--- 345
Db 1808 LSNPGNVTSONIIFTDVLPEGITFISGTLTNDSGTQIGNPATGIQIGNPGSTATIVI 1867
Qy 346 NPVTN--SVKP-----NSQQFV-----PHLSITL-----DENV 373
Db 1868 NALVTNIPSNIPSNFSSVQFAHVVDPSQPSVSTNLSNTVSTTTKSAITLTTKSGADKSV 1927
Qy 374 SSGDYIGTIOWTSPPSDSG---GANTNFWKIPDYGSSLAELASQAFAPVYPPGFNEVIV 430
Db 1928 ISVGD---TITYTTTITNTGNTAAAIKE-----TSAIPA-----NTTFIP 1965
Qy 431 FMASIPGPNOSG-SPNL---VPCLLPQEVIT-----HFISEQAPIQGEAALLHVDPD 479
Db 1966 NSVTINGVQSGVQPALGVNIPNIAPGETVTVTFQVNLVSFPSSSSIMGNDTILY----- 2020

QY 480 TNRNLGEFKLYPGGYLTCVPNSSTGPGQOLP-LDGVFVFASWVR 523
Db 2021 -----SYTDPNG--TPVTSTSTNVTNVPVLDAITMVKSDQ 2057

RESULT 8

US-10-282-122A-46893
Sequence 46893, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
PRIORITY FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 46893
LENGTH: 1185
TYPE: PR
ORGANISM: Bacillus anthracis
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (115)..(115)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (143)..(143)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (151)..(151)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (254)..(254)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (317)..(317)
OTHER INFORMATION: X=any amino acid

FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (369)..(369)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (466)..(466)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (470)..(470)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (619)..(619)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (750)..(750)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (808)..(808)
OTHER INFORMATION: X=any amino acid
US-10-282-122A-46893
Query Match 4.4%; Score 126.5; DB 12; Length 1185;
Best Local Similarity 20.9%; Pred. No. 0.077;
Matches 126; Conservative 77; Mismatches 223; Indels 177; Gaps 30;
QY 22 PVEVNT--ADPIP--IDPVAGS---STALATAGOVNLDIPWIIINNFVQAPQGEFTISPN 73
Db 177 IPANNVIFSDPIPSWTQFVAGSVIVDGTPLPSAS-----ITSGIGINTIIPN 223
QY 74 NTPGDVLEDLQ---GPHNLNPLSHLSOM---YNCWVGNMVRVVLGNAPFTAGKVIICC 127
Db 224 QTV-TIIFQVQIVSNPPTFTPELQNLAFVNFXYN-----VGNALQA----- 263
QY 128 VPPGFQSRSLTAQATLFPHVIAVRLDPEVPLED--VRNVLYHNNDTQPTMRLLCML 185
Db 264 -QPG--NVETNVFVTAIHSAILSAVKTASTAFANIGDITTYTLVLIQNRHFTLGNX--- 317
QY 186 YTPLRTGASGCT---DSFVVAGRVLTCPGPDFN----- 216
Db 318 ---ISQGIPIGCTTEVENSFAVNGN--TIPGANPNSGVNIPTVSAGSSLVTFLSRXLHS 372
QY 217 -----FLFLVPPTEVEKTRPFTVPNIPKYLNSRIPNPIEG----- 253
Db 373 TPPSNPITNVASIQFAFIVDPAAPPVGTGT--SNSASTQINNATVTTLEADRTIVSIGD 431
QY 254 -MSLSPDOTQNVQF--QNGRCTIDGOLGTPVVSQOLCKFRGRTISGORVNLNLTLDGSP 311
Db 432 IITYTATLTNTGNFPANSVLLINGVPEGALFVPIX--CHAXRDTFSEHQV-----QLSGIP 485
QY 312 -----FMAFAAPAPAGF---PDLGSCDWHIEMSKIIPNSSTONNPIVTVNSV 353
Db 486 VGI IAPGNSATITTFQFLANSIP--PQAIINQALTSYTVIVDPSQPPVATSSNTVTTAV 544
QY 354 KPNQQFVPHLSSITLDENVSSGGDIYIGTIQWTSPPSDSGGANTFWKIPDYGSSLAELAS 413
Db 545 VDASLSVKNKTDSDIVQSTD-----GTITYTVVIQNNGNTANTVTLTDL---VPEGT 593
QY 414 OLAPAVYPPGENEVIVYFMASIPGNQSGPNLVPCLLPQEVITHFISEAPIQGEAALL 473
Db 594 ALIP-----NSVTINSI-SIPGGRSKGKSHKXTSLAPSEIVPGHI PKLSILLFQSG-- 643
QY 474 HYVDPDTNRNLGEPKLYPGGYLTCVPNSSTGPGQOLP-LDGVFVFASWVR 533
Db 644 -----NPIFNARY----GLYFYCGSNCSYCLEQL-----LRNQLSHKQFMRVFL 687
QY 534 GPA 536
Db 688 KPS 690

Db 366 VPTGTGSAENVVLRDSIPN-----GTFVAGSVTVGGVTPQNPANPATGINL 412
QY 279 GTTPVSVSQLCKRGRIITSGQRVLNLTLDGSPFMAFAPAPAGFPDLGSCDWHIEMSKI 338
Db 413 GTIPNNTQIRIVTFQVRITS-----FPNPNP-----I 438
QY 339 PMSSTQN-----NPVITNSVKPNSQOFVPHLSSITLDENVSSGDDYIGTIQWTPPP 389
Db 439 PNRAMVSYQFRPRFVGSPLITSSSSNTVQTTVQATISMQKSDVLQATLNDVLTVTNVV 498
QY 390 SDGGANTNFKIPDYGSSLAELASQALAPAVPPGNEVIVYPMASIPGNSQSGSNLYPC 449
Db 499 TNGNVTAN--NVIFVDSIPAGTTFVA-----NSVIVNGVAR--PGANPASSINLGI 547
QY 450 LLPQEVITHFISEQAPIQGEAALLHYVDPTNRNLGEEKLYPGGYLTCVPSNSSSTGPOQL 509
Db 548 NASQTTVVR-----QVRVTSNPLNPIPNRASATFTFP-----VPGQQ 587
QY 510 PLDG 513
Db 588 PVSG 591

RESULT 10
US-10-282-122A-67793
; Sequence 67793, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67793
; LENGTH: 6310
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-10-282-122A-67793

US-10-282-122A-45763
; Sequence 45763, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45763
; LENGTH: 2358
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-10-282-122A-45763

Query Match 4.3%; Score 125.5; DB 12; Length 2358;
Best Local Similarity 22.1%; Pred. No. 0.27;
Matches 120; Conservative 57; Mismatches 174; Indels 193; Gaps 28;

QY 18 ACQVPEVNTADPID--PVAGSSTAL--ATAGVNLDPMLNMFVQAPQEFISPN 74
Db 193 AGQ---EITDASSPFVDATISGFATPATGAVTGRV-----LVTAQEGDSNIS--- 236
QY 75 TPGDVLFDLQGLPHLNPFLS-----HLSQMYNGWGNMVRVVLGNAFTAGKV 123
Db 237 --GD-----QURFGNANTTVALGPRPNANFFQSQCND-----SGNLDTSGF 280
QY 124 IICCVPPGQSRTLSAQNTLPHVJADVRLDPEVPLEDVNV-----LYHNNDTQPTM 179
Db 281 -----GDLNQLPLGVALA-----VARQGWDTNVDASSSLVNNQTSATV 318
QY 180 RLICMLYTLPLRTGASGGTDSFVVA--GRVLTCPGPDFNLFVLPPTV-----EQTRPFT 233
Db 319 RFV-----TNG-----DGYAAGFGVQIDATGPIINPVKSVNRTVAGVGDTLTYIT 365
QY 234 VPNIPKLYSN-----SRIPNPIEGMSLSPDQTONVQFQNGRCTIDG--QP-----L 278

Query Match 4.3%; Score 123.5; DB 12; Length 6310;
Best Local Similarity 21.4%; Pred. No. 1.9;
Matches 128; Conservative 56; Mismatches 204; Indels 211; Gaps 29;

QY 4 ASKATPESADGATG---AGQLVPEVNTADPIPDPVAGSSTALATAGOVNLDPIWNN 59
DB 3576 APSATFTADDSTAPAPAGLITPGGSA-----IQSGGAGSTV-EVKLAGTGLVGT 3627

QY 60 FVQAPQGEFTI--SPNNTPGDVLFDLQGLPHLNPFLSHLSMYNGVGNMR-----VR 110
DB 3628 VVVPAGGSTVPLSPAQLDQQA-----LNVLTDAAGNISQSQIPAPDITPALPTDVA 3682

QY 111 VVLAGNAFT-----AGKVIICCV-PPGFQSRKTSIAQ--ATLPHVIVADVR 153
DB 3683 VSSDGTAVTGNAPGASSVTVSDAGNVITVAVNPDPGSFVPLDTPQNNQTVTVVVTDA 3742

QY 154 TLD--PVEVPLEDVRNVLHNNDTQPTMELLCMLVTLPLTGASGCT-----DS 200
DB 3743 GNDAPSVSTADPTTN-----PEPATGL-----TVSPDGSTVGTAEPGSTVEVRNP 3790

QY 201 FVAGRVLTCPGDFNFLE-LVPPTVEQKTRPFTV-----PNIPLKYLNSRIPNP- 250
DB 3791 DIVRGVTA--GPDGTFVIVAPPLASGETVDVVIDPAGNESPELPLTGPTGTEVATPS 3848

QY 251 -----IEGMSLS-----PDQTNVQFQN-----GR 270
DB 3849 ALAISVDGELLTQGTGVSGLITVTSGTTLGSAATGSDGTFVFFQNAQLQVSAK 3908

QY 271 CTIDGCP-----LGTFP 282
DB 3909 ATVDGQPSVPAIIVANDTTAPDPTQVNLNATGSLTGQGEVATVRVTDLOGTLTGAT 3968

QY 283 VGSVOL--CKFRGRTSGORVNLNLTGDSPFMAFAAPAPAGFPDLGSCDWHIENSKI 340
DB 3969 VDSNGLFSVSPAVANGNLI-VTQADAGNVSLAATLQA--PDLQA-----4013

QY 341 SSTONNPIVTSVKPNSQOQFPHLS-----SITLDENVSSGG-----DYIGTIQWT 386
DB 4014 -----PLAASNLNLSAATV--LSQGEAGASVTVRD--ASGAILATGTVNSQGFQIT 4063

QY 387 SPDSGSGGANTFWKIPDYGSSLAERASQAPAVYPPG--FNEVIVVFMAISIPGNQSGS 443
DB 4064 LPSAQVTGSLQVTLSDAAGNVSPASLATPDHTPPAAISNPVLSQDGRQLSGSGEAGA 4122

RESULT 11
US-10-694-874-3
; Sequence 3, Application US/10694874
; Publication No. US20040097713A1
; GENERAL INFORMATION:
; APPLICANT: CELL SIGNALING TECHNOLOGY, INC.
; APPLICANT: POLAKIEWICZ, Roberto
; APPLICANT: LI, Yu
; APPLICANT: WU, Jiong
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR PHOSPHORYLATED IRS-1/2 (Ser1101/Ser1149)
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CST-209
; CURRENT APPLICATION NUMBER: US/10/694,874
; CURRENT FILING DATE: 2003-10-28
; PRIOR APPLICATION NUMBER: US 60/422,409
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1231
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-694-874-3

Query Match 4.1%; Score 119.5; DB 16; Length 1231;
Best Local Similarity 20.5%; Pred. No. 0.35;
Matches 98; Conservative 68; Mismatches 140; Indels 171; Gaps 29;

QY 156 DPVEVPLEDVRNVLHNNDTQPTMELLCMLVTLPLRTGASGCTDSFVWAGRVLTCPGPDF 215
DB 273 NPISVPLR-----RHLLNPPPSQ--VGLTRRSRTESITATSPASMVGGK-----PGS-- 318

QY 216 NFLFLVPPVTEOK---TSPFTVPNIPL-----KVLNSRIPNPIEGMSLSPDQTO 262
DB 319 ---FVRASSDGEGTMSRPASVDGSPVSPFTNRTAHRHRGSSRLHPPL-----NH8R 368

QY 263 NVQFQNGRGTIDQPLGTTTPVSVSQL-----CKFRGRTISGQORVNLNLTGDSPP 312
DB 369 SIPMSSRCS-----PSATSPVLSLSSSTSGHGSTDCLPFRSSA-----SVSGS-- 414

QY 313 MAFAPAPAGP---PDLGS--CDWHIEMSKIPINSTONNPIVTSVKPNSQOQFVPHLSI 367
DB 415 ---PSDGGFTISSDEYGGSPCDPRSPFRSVTPDLSLGHTPPARGBEELSNIYICMGKGAS 469

QY 368 TLDEN-----VSSGCD---YI-GTIQWTS-PSD--SGGAN-----T 397
DB 470 TLAAPNGHYILSRGGNGHRYIPGANLGTSPALPGDEAAGAADLDNRKTRTHSAGTSP 529

QY 398 NFWKLPDYS--SLAASQAPAVYPPGFNEVIVVFMAISIPGNQSGSNLVPCLLPQBY 455
DB 530 SHQKTPSQSSVASIEEYTEMPPAAYPPG-----GGSGRLPGYRHSAP 572

QY 456 I-THFISQAPIQGEAALLHYVD-----PDNTRNL-----GEFKLYP-----491
DB 573 VPTHSYPE-----EGLEMHLERRGGHHRPDTN-PLMTDDGYMPMSPGAPVPSNRKGN 625

QY 492 GYLTCVNSST-----GQQLPDGVFVFWASVSRFYQLKPVGTAGPARG 538
DB 626 GDMPSPKSVSAPOQIINPIRRHPQVDPNG-----YMMSPSGSCSPDIG 672

RESULT 12
US-10-055-569A-43
; Sequence 43, Application US/10055569A
; Publication No. US20040024181A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli, Raha A
; APPLICANT: Spyttek, Kimberly A
; APPLICANT: Gilbert, Jennifer
; APPLICANT: Casman, Stacie
; APPLICANT: Blalock, Angela
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Shenoy, Suresh
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit
; APPLICANT: Malyanker, Uriel
; APPLICANT: Stone, David
; APPLICANT: Millet, Isabelle
; APPLICANT: Smithson, Glenda
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Taupier Jr., Raymond J
; APPLICANT: Anderson, David W
; TITLE OF INVENTION: No. US20040024181A1 Human Proteins, Polynucleotides Encoding Th
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 21402-191
; CURRENT APPLICATION NUMBER: US/10/055,569A
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: 60/243,642
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,320
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,592
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,681
; PRIOR FILING DATE: 2000-10-27

PRIOR APPLICATION NUMBER: JP 2001-246789
PRIOR FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 2070
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1254
LENGTH: 1564
TYPE: PRT
ORGANISM: Homo sapiens
US-10-292-798-1254

Query Match 4.0%; Score 117; DB 15; Length 1564;
Best Local Similarity 20.8%; Pred. No. 0.87;
Matches 115; Conservative 50; Mismatches 181; Indels 208; Gaps 27;

Qy 10 PSADGATGAGQVPEVNTADPIDVAGSSTALATAGQVNLIDPWIINNFPVQAFQGEFT 69
Db 1160 PHARASAGICSLVPS-----VCRSPASLAVAALGSGV-----CHGGAFS 1198

Qy 70 ISFNNTPGDVLFDLQ---LQPHLNPFLSHLSQMYNGWGNMVRVVLGNATAGKVIIC 126
Db 1199 CS-----LVDCQGRGDGAGARELLSLPQGGSGYGN-----LVP 1233

Qy 127 CVPPGFSRTLSIAQATLPHVIADYRTLDPEVEPLEDVRNVLXHNNDTQPTMRLLCMLY 186
Db 1234 QSPPPVPSRTVPSVTSSVPTVPSVPS-SSCAVPHSPHPLH-----LCPLH 1281

Qy 187 TPLRTGASGCTDSFVAGRVLTCPGDFNLFVPTVEQKTRPFTVNIPLKYLNSR 246
Db 1282 SPLSP-----LHLLCPPTVPSVPSVSCVLPQSP-----QS 1310

Qy 247 IPNPIEGMS-----LSPDQT-QNVQFQNGRCTIDGQPLGTTTPSVSOLCKPGRITSQRV 301
Db 1311 SPPVPSQSSPPVPSPPSQSSPPVSSSTVPSVSSCPPTVPS-----VTSSCA 1363

Qy 302 LNLIELDGSFMAFAAPAGFPDLGSCDWHIEMSKI-----PNSSTQNNP----- 347
Db 1364 L-----PHSPLHHP-LLLCPHSPHSLHLLCPPAQSPQSPPPVPSRTVPS 1408

Qy 348 IVTNSVKPNSQQFVPHLSITLDENVSSGGDIQTIQWTSPPSDSGANTFWMKIPDYGS 407
Db 1409 VTSSSVPPSQSPPPSPPL-----SPPS-----PVPSPQS 1439

Qy 408 SLAELASQLAPAVYPGFEVIVYFMASIPGPNQSGS-----ENLVP-----CL 450
Db 1440 PLSPPPVSPCI-PP-----VTSFSCASPPSQSPQSPPLPPVPHPLLPNAPHRGAVAL 1493

Qy 451 LPQEIYTHFISQ---APIQGEAALLHYVDPTN--RNIGFKLYPGGYLTCTVFNSS----- 502
Db 1494 LPAPHGASKLHQDRLPGGRSELLQWVLP IQHPCHARGE-----PG---TCCPGQGNHV 1546

Qy 503 -----STGPOQ 508
Db 1547 QWPPRAALS LGPGQ 1560

RESULT 14
US-10-369-493-2294
Sequence 2294, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374

PRIOR APPLICATION NUMBER: 60/243,863
PRIOR FILING DATE: 2000-10-27
PRIOR APPLICATION NUMBER: 60/244,443
PRIOR FILING DATE: 2000-10-31
PRIOR APPLICATION NUMBER: 60/245,029
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/244,995
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/245,293
PRIOR FILING DATE: 2000-11-02
PRIOR APPLICATION NUMBER: 60/245,315
PRIOR FILING DATE: 2000-11-02
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 137
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 43
LENGTH: 609
TYPE: PRT
ORGANISM: Xenopus laevis
US-10-055-569A-43

Query Match 4.1%; Score 117.5; DB 16; Length 609;
Best Local Similarity 20.8%; Pred. No. 0.18;
Matches 96; Conservative 44; Mismatches 156; Indels 165; Gaps 24;

Qy 23 PEVNTADP-----IPIDPVAGSSTALATAGQVNLIDPW-----IINNFPVQAFQGEFTISP 72
Db 265 PVAPTADPWGASAAPADPWAGGATPASVFAAAAADPPWGGPPVATGSSSDPWGT-GVQT 323

Qy 73 NNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNATAGKVIICVVP--P 130
Db 324 NSTPGD-----PWGDT-----QAVTSADVKSVDSPWNP 351

Qy 131 GFQSRITLSIAQA--TLPHVIADV-RTLDPVEVEPLEDVRNVLXHNNDTQPTMRLLCMLYT 187
Db 352 GGSQATTAIPSDPWSSPPVAQSVKKAADPWAPASFSDP--WGGSPSKP----- 400

Qy 188 PLRTGGASGCTDSFVAGRVLTCPGDFNLFVPTVEQKTRPFTVNIPLKYLNSRI 247
Db 401 --NTNGTGMELD--LLAGEV-----PMSRSLGSKS 426

Qy 248 NPTEGMSLSPDQTQNVQFQNGRCTIDGQPLGTTTPSVSOLCKPGRITSGQVNLNLT 307
Db 427 PDADFDMSTMS-----GSLCDPS-----NPTRK 448

Qy 308 DGSFPMFAFAAPAGFPDLGSCDWHIEMSKIPNSSTON-----NPIVTNSVKPNSQQ 359
Db 449 TPESFLG---PNAALVDLDSL---ISKSTLQNTKTSNPFVLTGTPTATNPFPQPNQ- 500

Qy 360 FVPHLSITLDENVSSGGDIQTIQWTSPPSDSGANTFWMKIPDYGSSLAELASQLAPAV 419
Db 501 -----SSLTNQLRSSVMTLG-----QVTPAGQTPAT---IP-FASPMSVSPMAGI 546

Qy 420 YPPGFNEVI-VYFMASIP-GPNQSGPNLV--PCILLPOEYI 456
Db 547 PLANNAPVWGMQPMAGVPVGTILAPGVGWLPPMMPQQLV 587

RESULT 13
US-10-292-798-1254
Sequence 1254, Application US/10292798
Publication No. US20030235833A1
GENERAL INFORMATION:
APPLICANT: SUWA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABURATANI, HIROYUKI
TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
FILE REFERENCE: 084335/166
CURRENT APPLICATION NUMBER: US/10/292,798
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: 10/017,161
PRIOR FILING DATE: 2001-12-18

SEQ ID NO 2294
LENGTH: 913

TYPE: PRT

ORGANISM: Schizosaccharomyces pombe

FEATURE:

NAME/KEY: unsure

LOCATION: (1)-(913)

OTHER INFORMATION: unsure at all Xaa locations

US-10-369-493-2294

Query Match 4.0%; Score 114.5; DB 15; Length 913;
Best Local Similarity 19.8%; Pred. No. 0.64;
Matches 117; Conservative 96; Mismatches 212; Indels 165; Gaps 32;

QY 9 TPSAGATGA--GQLVPEVNTADPID-----PVAGSSTA-----LATAG--QVN 50
Db 243 TVPAVGSTTSTVPDVVXAGNVEVTEPIDMVSIVTPAVGSTTSTVPDVPXAGIVEVN 302
QY 51 LIDPWLNNFVQAPQGEFTI--SPNNTPGDVLGLPHLNPFLS--HLSQYNGWGNMR 108
Db 303 GIDPETIVSVTPAVGSTTMTAPDEVXAVIVDIK---ETDPEITVSVTPAVGSTTSTV 359
QY 109 VRVW-LAGNAF---TAGKVIICVPPGFQSRILSIAQATLFPHVI-----ADVRLDPRVE 159
Db 360 TRVPLAGNVEVYPTETPMVSVIVPAVGSTTSTV-----PEVPLAGNVEVKULEPET 413
QY 160 -----VPLEDVRNVLVHNNDTQPTMRLLCMLYTPRLTGASGGTDS 200
Db 414 MVSIVTPAVGSTTSTVPEVVPVPLAGNVEVNALEPDT-----MVSIVVP-----AVGSTTS 462
QY 201 FV-----VAGRVLTCPGDPFNFLPLVPTVEQKTRPFTVPNPLKYLNSRIPNPIEG-M 254
Db 463 TVPEVVPVPLAGNVEVYPTETPMVSVIVTPAVGSTTSTV-----TVPEV-----PLAGNV 509
QY 255 SLSPDOTQNVQFNGRCTIDGQPLGTPVSVLSQCKFRGRI-----TSQORVNLNLTLDGS 310
Db 510 EVYPTETPM-----VSVTVPAVGSTTSTVPEVVPVPLAGNVEVYPTETPMVSVI----- 558
QY 311 PFMAFAAPAGAPFDLGSCDWHIEMSKIPNSSTQNNPIVITNSVKPNSSQPV 367
Db 559 -----VPAVGSTN-----STVPEVVPVPLAGNVEVNALEPDTMVSIVVPAVGSTT 601
QY 368 -TLDENVSSGGDYIGTIQWTSPPSDGGANTFWKIPDYGSSLAESAQLAP-----AVYP 421
Db 602 STVPEVVPVPLAGNV--EVKLEPETMVSIV-----VPAVGSTTSTVPEVVPVPLAGNVEVYP 653
QY 422 PGFNEVIVYFMASIRPGNOSGSPNLVP-----CLLPQEVY-----THFISEQ 463
Db 654 TE-PETMVSIVVPAVGSTTSTVPEVVPVPLAGNVEVKVLEPETMVFVIVPAVGSTTSTVPEV 712
QY 464 APIQGEAALLHYVDPDTNRNLGFEKLYPGGYLTCPVNSSTGPPQQLPLDG 513
Db 713 VPLAGNVE-VNALEPDTMVSIV-----IVPAVGSTTSTVPEVVPVPLAG 752

RESULT 15

US-10-618-581-5

Sequence 5, Application US/10618581

Publication No. US20040077524A1

GENERAL INFORMATION:

APPLICANT: MANNING, GERARD

APPLICANT: SUDARSANAM, SUCHA

TITLE OF INVENTION: METHOD FOR SELECTIVELY INHIBITING FUNGAL GROWTH

FILE REFERENCE: 034536/0323

CURRENT APPLICATION NUMBER: US/10/618,581

CURRENT FILING DATE: 2003-07-15

PRIOR APPLICATION NUMBER: 60/395,624

PRIOR FILING DATE: 2002-07-15

NUMBER OF SEQ ID NOS: 42

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 5

LENGTH: 1212

TYPE: PRT

ORGANISM: Candida albicans

US-10-618-581-5

Query Match 4.0%; Score 114.5; DB 16; Length 1212;
Best Local Similarity 20.4%; Pred. No. 0.99;
Matches 114; Conservative 72; Mismatches 215; Indels 157; Gaps 29;

QY 23 PEVN--TADPIDPVAGSSTALATAGQVNLIDPWIINNFVQAPQGEFTISN----- 73
Db 591 PEVSQPLPEPKFVPPPEIINPAVATQAQANMTAPKIVETSETPQR--TLDSKQSVDBK 648
QY 74 -NTPGDVLPDLQLGPHLNPFLSHLSQYNGWGNMRVVRVVLGNAGTACKVIICCVPPGF 132
Db 649 PSAPGP-----STAVPEQAHTTS-----VPSSP 671
QY 133 QSRTLSIAQATL-FPHV-----IADVRLDPRVEPLEDVRNVLVHNNDTQPTMRLLCMLYT 187
Db 672 VKTQTSIDEDQLSIFEQSQSPRTSTPTQLDPAKV----- 704
QY 188 PLRTGASGGTDSFVAVAGRVLTCPGDPFNFLPLVPTVEQKTRPFTVPNPLKYLNSRI 247
Db 705 ---VGGSSGSAISAPNAG-----SCAGFNSLLR-----RLSSKVKYGASSPKR--STSPS 749
QY 248 NPPIEGMSLSPDOTQNVQFNG---RCIIDGQPLGTPVSVLSQCKFR---GRITSGQRV 301
Db 750 PN-VEGLSPQPTKA--DPMVVRGVSMKVTAKEKQTNTRPPK-SELIKKQPHGRSSS----- 802
QY 302 LNLTELDSPPFMAFAAPAGAPFDLGSCDWHIEMSKIPNSSTQNNPIVITNSVKPNSSQPV 361
Db 803 -TSNKKQGFIPVEYLP-----LPTIDTNTVIVSDGAKQQLT-V 841
QY 362 P-----HLSITLDENVSSG---GDYIGTIQWTS-----PPSDGGANTFWKII--PDYGS 408
Db 842 PSTARHMHPTARAKSVGGGHMRKDSYGRVSHGSQNPPLPPLPTSMASQNSQEVVVKDTSBG 901
QY 409 LAEASQALAPAVYP--PGFNEVIVYFMASIRPGNOSGSPNLVPCLLPOEVIHIFISEQAPI 466
Db 902 PFDDVQLDVGYQEVLPQLTESIIEQYNISKEN--SMPSIEHC--KTLFLKGFPSVQTTS 957
QY 467 QGEAALLHYVDPDTNRNLG-EFKLYPGGYLTVCV-----PNSSSSTGPPQQLPLDGVPVFAS 519
Db 958 AKPLPVIRYININVLKGVKQFVKGGFV-CMHTPSVQPSVQPSHNSDELDEENKLYGDAPFASK 1016
QY 520 WVSRYQLKPVGTAGPAR 537
Db 1017 SSDSPEAAEPEGSKTFSR 1034

Search completed: June 1, 2004, 14:04:33

Job time : 36.7758 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:36:02 ; Search time 10.4966 Seconds
(without alignments)
4985.230 Million cell updates/sec

Title: US-09-926-799-4
Perfect score: 2891
Sequence: 1 MMASKDATPSADGATGAGQ.....YOLKPVGTAGPARGRLGVR 544

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2234.5	77.3	546	B37491	major capsid prote
2	2096	72.5	530	B37471	capsid protein - N
3	1171.5	40.5	542	S60616	capsid protein - h
4	1154	39.9	539	S40111	capsid protein - h
5	347.5	12.0	576	A53982	capsid protein - E
6	304	10.5	2344	RRWRH	genome polyprotein
7	298	10.3	2344	S53399	genome polyprotein
8	297	10.3	2344	S64740	genome polyprotein
9	285	9.9	703	C48562	coat protein - San
10	281	9.7	702	A48562	coat protein - San
11	274	9.5	668	VCWVFC	coat protein - fel
12	272.5	9.4	668	JQ2354	capsid protein - f
13	264.5	9.1	671	VCWVFF	coat protein - fel
14	263.5	9.1	668	VCWVFF	coat protein - fel
15	253	8.8	668	JQ2356	capsid protein - f
16	148	5.1	2194	GNNE7	genome polyprotein
17	141	4.9	2508	S61441	surface-associated
18	137	4.7	2206	S03822	genome polyprotein
19	136.5	4.7	3473	A46112	genome polyprotein
20	136.5	4.7	3473	S27927	polyprotein - rice
21	135	4.7	2206	GNNY4P	genome polyprotein
22	134	4.6	2175	GNNYBE	genome polyprotein
23	132.5	4.6	2206	S3327	genome polyprotein
24	123	4.3	915	S36327	clathrin assembly
25	122	4.2	2205	GNNY2W	genome polyprotein
26	122	4.2	2207	GNNY5P	genome polyprotein
27	121	4.2	2183	T42764	coagulation factor
28	120.5	4.2	896	S63326	clathrin assembly
29	119.5	4.1	1231	S30185	insulin receptor s

30	117.5	4.1	1235	1	S16948	insulin receptor s
31	116.5	4.0	2207	2	S09553	genome polyprotein
32	116	4.0	3016	2	S77300	hypothetical prote
33	115.5	4.0	929	2	A44048	genome polyprotein
34	115.5	4.0	1367	2	T33819	hypothetical prote
35	115.5	4.0	2179	1	GNNYH4	genome polyprotein
36	113.5	3.9	745	2	G72453	hypothetical prote
37	113.5	3.9	799	2	S75351	hypothetical prote
38	113.5	3.9	2333	1	GNNY2F	genome polyprotein
39	113.5	3.9	3085	2	T00327	polyprotein - infe
40	113	3.9	1075	2	A57377	transcription fact
41	113	3.9	2150	1	GNNYH2	genome polyprotein
42	113	3.9	2206	1	GNNY21	genome polyprotein
43	112.5	3.9	537	2	T04822	hypothetical prote
44	112.5	3.9	648	2	PC4395	mucin 3 - human (f
45	112	3.9	554	2	T01240	laccase (EC 1.10.3

ALIGNMENTS

RESULT 1

B37491
- major capsid protein [similarity] - Southampton virus
N:Alternate names: orf2 protein
C:Species: Southampton virus
C:Date: 22-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 18-Aug-2000
C:Accession: B37491
R:Lambden, P. R. *Journal of Virology* 67:1993

Sequence 259, 516-519-1993
A:Title: Sequence and genome organization of a human small round-structured (Norwalk-like) virus
A:Reference number: A37491; MUID:93142023; PMID:8380940
A:Accession: B37491
A:Status: not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-546 <I>AMs
A:CROSS-references: GB:L07418; NID:gl236787; PIDN:AAA92984.1; PID:g295114
A:Note: sequence extracted from NCBI backbone (NCBIP:123458)
A:Note: small round-structured virus, SRSV, Norwalk virus, Norwalk-like virus, serotype
C:Superfamily: human calicivirus capsid protein
C:Keywords: glycoprotein
F:303,340,441/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 77.3%; Score 2234.5; DB 2; Length 546;

Best Local Similarity 74.9%; Pred. No. 1e-149;

Matches 408; Conservative 58; Mismatches 78; Indels 1; Gaps 1;

QY 1 MMASKDATPSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGQVNLDPVNNP 60

DB 1 MMASKDAPQSDAGSAGQLVPEVNTADPLPEPVAGPTTAVATAGQVNMIDPVNNP 60

QY 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNRVVVLGNAFTA 120

DB 61 VQSPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGNRVVVLGNAFTA 120

QY 121 KRVIICCVPPGQSQTLSIAQATLFPPIADVRLTDPVEVPLEVRNVLVHNDTQPTMR 180

DB 121 KRVIICCVPPGQSQTLSIAQATLFPPIADVRLTDPVEVPLEVRNVLVHNDTQPTMR 180

QY 181 LLCMLYTLRTGGAGGTTDSFVAGRVLTCPGDFNFLELVPTVEQKTRPTVNIPLK 240

DB 181 LLCMLYTLRTGGAGGTTDSFVAGRVLTCPGDFNFLELVPTVEQKTRPTVNIPLK 240

QY 241 YLSNRIINPITGMSLSPDQTONVQFNGRCRTIDGQPLTTPVSVSQLCKPGRITTSQGR 300

DB 241 YLSNRIINPITGMSLSPDQTONVQFNGRCRTIDGQPLTTPVSVSQLCKPGRITTSQGR 300

QY 301 VLNLTELDSPPMAFAPAPAGFPDLGSCDWHIEMSKIPNSSTONNPIVNSVKPNSQQF 360

DB 301 VLNLTELDSPPMAFAPAPAGFPDLGSCDWHIEMSKIPNSSTONNPIVNSVKPNSQQF 360

QY 361 VPHLSITLIDENVS-SGGDYIGTITQWTSPDSDGGANNTFWKIPDYGSSLABASQAPAV 419

DB 361 VPHLSITLIDENVS-SGGDYIGTITQWTSPDSDGGANNTFWKIPDYGSSLABASQAPAV 419

361 VPHLGSIOFDEVFNHTGDICTIEWISQSPSTPPCTDINLWEIPDYGSSLSQANLAPPV 420
 420 YPPGNEVLVYFMASIPGNSQSNLVPCLLPQYIYTHFISEQAPIQGEAALLHYVDPD 479
 421 FPPGGEALVYFVSAPFNNRNSANDVPCLLPQYIYTHFISEQAPTMGDAALLHYVDPD 480
 480 TNRNLGEFLKYPGGYLTCVPSNSSSTGPOQLPLDGVFVFSWVSRYQKPVGTAGPARGR 539
 481 TNRNLGEFLKYPGGYLTCVPGNGVAGPQQLPLNGVFLFVSWVSRYQKPVGTASTARGR 540
 540 LGVRR 544
 541 LGVRR 545

RESULT 2
 B37471
 capsid protein - Norwalk virus
 C:Species: Norwalk virus
 C:Date: 24-Feb-1994 #sequence_revision 25-Apr-1997 #text_change 28-Jul-2000
 C:Accession: B37471
 R:Jiang, X.; Wang, M.; Wang, K.; Estes, M.K.
 Virology 195, 51-61, 1993
 A:Title: Sequence and genomic organization of Norwalk virus.
 A:Reference number: A37471; MUID:93303939; PMID:8391187
 A:Accession: B37471
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: genomic RNA
 A:Residues: 1-530 <JFA>
 A:Cross-references: GB:M87661; NID:g1061311; PIDN:AAB50466.1; PID:g1061313
 A:Note: sequence extracted from NCBI backbone (NCBIP:134157)
 C:Superfamily: human calicivirus capsid protein

Query Match 72.5%; Score 2096; DB 2; Length 530;
 Best Local Similarity 72.5%; Pred. No. 5.6e-140;
 Matches 395; Conservative 55; Mismatches 79; Indels 16; Gaps 8;

QY 1 MMASKDAPTSADGATGAGQLVPEVNTADIPIDPVAGSSSTALATAGQVNLDPWIINNF 60
 DB 1 MMASKDAPTSADGATGAGQLVPEVNTADIPIDPVAGSSSTALATAGQVNLDPWIINNF 60

QY 61 VOAPQGETTISPNNTPGDVLFDLQGLPHLPFLSHLSQMYNGWGNRVVWVLAGNAFTA 120
 DB 61 VOAPQGETTISPNNTPGDVLFDLQGLPHLPFLSHLSQMYNGWGNRVVWVLAGNAFTA 120

QY 121 GKVIICCPVPGFQSGRTLSIAQATLFPVHIAVRLDPEVPLEVDRVNLVHND-TQPTM 179
 DB 121 GKVIICCPVPGFQSGRTLSIAQATLFPVHIAVRLDPEVPLEVDRVNLVHND-TQPTM 179

QY 180 RLLCMLYPLRTGGASGTDGTFVAVAGVLTCPGDFNLFVLPVPTVEQKTRPFTVNIPL 239
 DB 180 RLLCMLYPLRTGGASGTDGTFVAVAGVLTCPGDFNLFVLPVPTVEQKTRPFTVNIPL 239

QY 240 KYLSNSRIPNPIEGMSLSPDQTONVQFONGRCTLGQPLGTPVSVSOLCKFRGRTISGQ 299
 DB 240 KYLSNSRIPNPIEGMSLSPDQTONVQFONGRCTLGQPLGTPVSVSOLCKFRGRTISGQ 299

QY 360 FVPHLSSTILDENVSSGGDYICTIOWTSPDSGGANTFNFKIPDYSSSLAEASQAPAV 419
 DB 360 FVPHLSSTILDENVSSGGDYICTIOWTSPDSGGANTFNFKIPDYSSSLAEASQAPAV 419

QY 420 YPPGFNEVYFMASIPGNSQSNLVPCLLPQYIYTHFISEQAPIQGEAALLHYVDPD 479
 DB 420 YPPGFNEVYFMASIPGNSQSNLVPCLLPQYIYTHFISEQAPIQGEAALLHYVDPD 479

QY 480 TNRNLGEFLKYPGGYLTCVPSNSSSTGPOQLPLDGVFVFSWVSRYQKPVGTAGPARGR 539
 DB 480 TNRNLGEFLKYPGGYLTCVPSNSSSTGPOQLPLDGVFVFSWVSRYQKPVGTAGPARGR 539

QY 490 YPPGFNEVYFMASIPGNSQSNLVPCLLPQYIYTHFISEQAPIQGEAALLHYVDPD 479
 DB 490 YPPGFNEVYFMASIPGNSQSNLVPCLLPQYIYTHFISEQAPIQGEAALLHYVDPD 479

QY 540 LGVRR 544
 DB 540 LGVRR 545

540 LGVRR 544
 542 LGVRR 530

RESULT 3
 S60616
 capsid protein - human calicivirus (strain Melksham)
 C:Species: human calicivirus
 C:Variety: strain Melksham
 C:Date: 23-May-1997 #sequence_revision 23-May-1997 #text_change 28-Jul-2000
 C:Accession: S60616
 R:Green, S.M.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
 Virus Res. 37, 271-283, 1995
 A:Title: Capsid diversity in small round-structured viruses: molecular characterization
 A:Reference number: S60615; MUID:96136658; PMID:8533462
 A:Accession: S60616
 A:Molecule type: genomic RNA
 A:Residues: 1-542 <GRE>
 A:Cross-references: EMBL:X81879; NID:g976077; PIDN:CAA57462.1; PID:g976079
 A:Experimental source: strain Melksham
 A:Note: it is uncertain whether Met-1 or Met-3 is the initiator
 C:Superfamily: human calicivirus capsid protein
 C:Keywords: capsid protein; coat protein

Query Match 40.5%; Score 1171.5; DB 2; Length 542;
 Best Local Similarity 45.6%; Pred. No. 7.8e-75;
 Matches 256; Conservative 87; Mismatches 176; Indels 43; Gaps 14;

QY 1 MMASKDAPTSADGATGAGQLVPEVNTADIPIDPVAGSSSTALATAGQVNLDPWIINNF 60
 DB 1 MMASKDAPTSADGATGAGQLVPEVNTADIPIDPVAGSSSTALATAGQVNLDPWIINNF 60

QY 61 VOAPQGETTISPNNTPGDVLFDLQGLPHLPFLSHLSQMYNGWGNRVVWVLAGNAFTA 120
 DB 61 VOAPQGETTISPNNTPGDVLFDLQGLPHLPFLSHLSQMYNGWGNRVVWVLAGNAFTA 120

QY 121 GKVIICCPVPGFQSGRTLSIAQATLFPVHIAVRLDPEVPLEVDRVNLVHND-TQPTM 179
 DB 121 GKVIICCPVPGFQSGRTLSIAQATLFPVHIAVRLDPEVPLEVDRVNLVHND-TQPTM 179

QY 180 RLLCMLYPLRTGGASGTDGTFVAVAGVLTCPGDFNLFVLPVPTVEQKTRPFTVNIPL 239
 DB 180 RLLCMLYPLRTGGASGTDGTFVAVAGVLTCPGDFNLFVLPVPTVEQKTRPFTVNIPL 239

QY 240 KYLSNSRIPNPIEGMSLSPDQTONVQFONGRCTLGQPLGTPVSVSOLCKFRGRTISGQ 297
 DB 240 KYLSNSRIPNPIEGMSLSPDQTONVQFONGRCTLGQPLGTPVSVSOLCKFRGRTISGQ 297

QY 360 FVPHLSSTILDENVSSGGDYICTIOWTSPDSGGANTFNFKIPDYSSSLAEASQAPAV 419
 DB 360 FVPHLSSTILDENVSSGGDYICTIOWTSPDSGGANTFNFKIPDYSSSLAEASQAPAV 419

QY 420 YPPGFNEVYFMASIPGNSQSNLVPCLLPQYIYTHFISEQAPIQGEAALLHYVDPD 479
 DB 420 YPPGFNEVYFMASIPGNSQSNLVPCLLPQYIYTHFISEQAPIQGEAALLHYVDPD 479

QY 480 TNRNLGEFLKYPGGYLTCVPSNSSSTGPOQLPLDGVFVFSWVSRYQKPVGTAGPARGR 539
 DB 480 TNRNLGEFLKYPGGYLTCVPSNSSSTGPOQLPLDGVFVFSWVSRYQKPVGTAGPARGR 539

QY 490 YPPGFNEVYFMASIPGNSQSNLVPCLLPQYIYTHFISEQAPIQGEAALLHYVDPD 479
 DB 490 YPPGFNEVYFMASIPGNSQSNLVPCLLPQYIYTHFISEQAPIQGEAALLHYVDPD 479

QY 540 LGVRR 544
 DB 540 LGVRR 545

RESULT 4
 S40111
 capsid protein - human calicivirus (strain Bristol isolate B493)


```
Db      1779 AGTATTASVPGTTTDMGDBGVATTSTVITAE:---NSSASIIATAGIGGPPQQVDOQETWRT 1835
QY      58  NNFOAQPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLVLAGNA 117
Db      1836 NFY---NDVFTWSVADAPGSLTYTVQHSPPQNPFATVLSQMYAGWAGMGRFRIVAGSG 1892
QY      118 FTAGKVIICCVPPGFO-SRTLSIAQATLFPFHVIADVRTLDPEVEPLEDVRNVLYH-NNDT 175
Db      1893 VFGGRVRAVIFPGIEIGPLEVRQ---FPHVVIDARSLEPTVITMPLRLPNMYHTPTGDP 1949
QY      176 OPTMRLLCMLYTPLTRGGASGDTDFVAVAGRVLTCTGPDFFNLFVLPVPTVEQKTRPFTVP 235
Db      1950 GLVPTLVLSVYNNL-INPFGSGTSAIQVT--VETRPSEDFEFVMI:----- 1991
QY      236 NIPLKYLNSRIPNPIEGMSLSPDQTNQVQFQNGRCTIDGQPLGTPV--SVSOLCKFRG 293
Db      1992 -----RAPSSKTVDISF:-----AGLLTTPVLTVGVGNDRNRNG 2024
QY      294 RITSGORVLNLTLDGSPFMAFAAPAGFPDLGSCDMHIEMSKI PNSTQNNPIVTNSV 353
Db      2025 QIVGLQ-----PVEGGF---STCNRHWNL-----NGSTYG----- 2051
QY      354 KPSNQFVPHLSITLDENVSSGDIYGIQTWSP-----PDSGGGANTN---F 399
Db      2052 -----WSSPRFGIDHRRGSASYSGSNATNVLOF 2080
QY      400 WKIPDYGSSIAE-ASQLAPAVYPFGFNEVIVFMASIP--GP-----NQS 441
Db      2081 W-YANAGSVDNPIISQVAPDGFDP-----MSFVPFNGPGIPAAQWVGFGAIWNSNS 2130
QY      442 GSPNLVCLLPQBYITHFISEQAPIQGEAALLHVDPDTNRNLGEFKLYPGGYLTICVPN- 500
Db      2131 GAVNVTTV---QAYELGF-ATGAP-----GNLQPTTNTS-----GAQTVAKSI 2169
QY      501 -SSSTGPGQPLDGVFVFASWV 521
Db      2170 YAVVTGTAQNP-GLFVWASGI 2190

RESULT 7
S55399 genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)
C:Species: rabbit hemorrhagic disease virus
A:Variety: isolate BS89
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 18-Jun-1999
C:Accession: S55399
R:Roosi, C.
submitted to the EMBL Data Library, May 1995
A:Reference number: S55399
A:Accession: S55399
A:Molecule type: genomic RNA
A:Residues: 1-2344 <ROS>
A:Cross-references: EMBL:X87607; NID:G854640; PIDN:CAA60910.1; PID:9854641
A:Experimental source: isolate BS89
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: polyprotein

Query Match 10.3%; Score 298; DB 2; Length 2344;
Best Local Similarity 24.9%; Pred. No. 2.1e-12;
Matches 140; Conservative 59; Mismatches 169; Indels 194; Gaps 28;

QY      4 ASKDATPSADGATGAGQLVPEVNTADPTIDPVAGSSPALATAG-----QVNLDIPWII 57
Db      1779 AGTATTASVPGTTTDMGDBGVATTSTVITAE:---NSSASIIATAGIGGPPQQVDOQETWRT 1835
QY      58  NNFOAQPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLVLAGNA 117
Db      1836 NFY---NDVFTWSVADAPGSLTYTVQHSPPQNPFATVLSQMYAGWAGMGRFRIVAGSG 1892
QY      118 FTAGKVIICCVPPGFO-SRTLSIAQATLFPFHVIADVRTLDPEVEPLEDVRNVLYH-NNDT 175
Db      1893 VFGGRVRAVIFPGIEIGPLEVRQ---FPHVVIDARSLEPTVITMPLRLPNMYHTPTGDP 1949
QY      176 OPTMRLLCMLYTPLTRGGASGDTDFVAVAGRVLTCTGPDFFNLFVLPVPTVEQKTRPFTVP 235
Db      1950 GLVPTLVLSVYNNL-INPFGSGTSAIQVT--VETRPSEDFEFVMI:----- 1991
QY      236 NIPLKYLNSRIPNPIEGMSLSPDQTNQVQFQNGRCTIDGQPLGTPV--SVSOLCKFRG 293
Db      1992 -----RAPSSKTVDISF:-----AGLLTTPVLTVGVGNDRNRNG 2024
QY      294 RITSGORVLNLTLDGSPFMAFAAPAGFPDLGSCDMHIEMSKI PNSTQNNPIVTNSV 353
Db      2025 QIVGLQ-----PVEGGF---STCNRHWNL-----NGSTYG----- 2051
QY      354 KPSNQFVPHLSITLDENVSSGDIYGIQTWSP-----PDSGGGANTN---F 399
Db      2052 -----WSSPRFGIDHRRGSASYSGSNATNVLOF 2080
QY      400 WKIPDYGSSIAE-ASQLAPAVYPFGFNEVIVFMASIP--GP-----NQS 441
Db      2081 W-YANAGSVDNPIISQVAPDGFDP-----MSFVPFNGPGIPAAQWVGFGAIWNSNS 2130
QY      442 GSPNLVCLLPQBYITHFISEQAPIQGEAALLHVDPDTNRNLGEFKLYPGGYLTICVPN- 500
Db      2131 GAVNVTTV---QAYELGF-ATGAP-----GNLQPTTNTS-----GAQTVAKSI 2169
QY      501 -SSSTGPGQPLDGVFVFASWV 521
Db      2170 YAVVTGTAQNP-GLFVWASGI 2190
```

```
QY      176 OPTMRLLCMLYTPLTRGGASGDTDFVAVAGRVLTCTGPDFFNLFVLPVPTVEQKTRPFTVP 235
Db      1950 GLVPTLVLSVYNNL-INPFGSGTSAIQVT--VETRPSEDFEFVMI:----- 1991
QY      236 NIPLKYLNSRIPNPIEGMSLSPDQTNQVQFQNGRCTIDGQPLGTPV--SVSOLCKFRG 293
Db      1992 -----RAPSSKTVDISF:-----AGLLTTPVLTVGVGNDRNRNG 2024
QY      294 RITSGORVLNLTLDGSPFMAFAAPAGFPDLGSCDMHIEMSKI PNSTQNNPIVTNSV 353
Db      2025 QIVGLQ-----PVEGGF---STCNRHWNL-----NGSTYG----- 2051
QY      354 KPSNQFVPHLSITLDENVSSGDIYGIQTWSP-----PDSGGGANTN---F 399
Db      2052 -----WSSPRFGIDHRRGSASYSGSNATNVLOF 2080
QY      400 WKIPDYGSSIAE-ASQLAPAVYPFGFNEVIVFMASIP--GP-----NQS 441
Db      2081 W-YANAGSVDNPIISQVAPDGFDP-----MSFVPFNGPGIPAAQWVGFGAIWNSNS 2130
QY      442 GSPNLVCLLPQBYITHFISEQAPIQGEAALLHVDPDTNRNLGEFKLYPGGYLTICVPN- 500
Db      2131 GAVNVTTV---QAYELGF-ATGAP-----GNLQPTTNTS-----GAQTVAKSI 2169
QY      501 -SSSTGPGQPLDGVFVFASWV 521
Db      2170 YAVVTGTAQNP-GLFVWASGI 2190

RESULT 8
S64740 genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)
N:Contains: VP60 protein
C:Species: rabbit hemorrhagic disease virus
A:Variety: isolate AST/89
C:Date: 12-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Jun-2000
C:Accession: S64740; S46944; S49018; S65012
R:Casais, R.; Martin-Alonso, J.; Boga, J.; Parra, F.
submitted to the EMBL Data Library, May 1995
A:Description: Genomic organization of rabbit hemorrhagic disease virus determined by d
A:Reference number: S64740
A:Accession: S64740
A:Molecule type: genomic RNA
A:Residues: 1-2344 <CAS>
A:Cross-references: EMBL:Z49271; NID:G1182032; PIDN:CAA89265.1; PID:G1150552
A:Experimental source: isolate AST/89
R:Boga, J.; Casais, R.; Marin, M.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra, F.
submitted to the EMBL Data Library, July 1993
A:Description: Molecular cloning, sequence and expression of the capeid protein gene fro
A:Reference number: S46944
A:Accession: S46944
A:Molecule type: genomic RNA
A:Residues: 1650-2344 <BOG>
A:Cross-references: EMBL:Z24757; NID:G515622; PIDN:CAA80881.1; PID:G515623
A:Experimental source: isolate AST/89
R:Parra, F.; Boga, J.A.; Marin, M.S.; Casais, R.
Virus Res. 27, 219-228, 1993
A:Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus supp
A:Reference number: S49018; MUID:93255896; PMID:8488721
A:Accession: S49018
A:Molecule type: genomic RNA
A:Residues: 1650-1796 <PAR>
A:Cross-references: EMBL:Z24757
A:Experimental source: isolate AST/89
A:Accession: S65012
A:Molecule type: protein
A:Residues: 1767-1779; 1875-1877, 'X', 1879-1881, 1936-1938, 'X', 1940-1941 <PAW>
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: blocked amino end; polyprotein

Query Match 10.3%; Score 297; DB 2; Length 2344;
Best Local Similarity 24.9%; Pred. No. 2.4e-12;
```

```
Matches 140; Conservative 58; Mismatches 170; Indels 194; Gaps 28;
QY 4 ASKADTPSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAG-----QVNLIDPMII 57
DB 1779 AGTATTASVGGTTDMDPGVVAITSVVTAEE---NSSASIAATAGIGGPPQVDQOETWRT 1835
QY 58 NNFOVAPQGEFTISPNPTGDLFDLQGLPHLNPFLSHLSQMYNGVGNMRVRVVLGNA 117
DB 1836 NFYY---NDVFTWSVADAPGSLYTVQHSPPQNPFTAVLSQMYAGWAGMQFRFVAGIG 1892
QY 118 FTAGVVICVPPGFQ-SRTLSIAQATLPPHHIADVRTLDPEVPLEDVRNVLVYH-NDNT 175
DB 1893 VFGRLVAIVPPGIEGPLEVRQ---PPHHVIDARSLPPTIYTPDLRPNMYHTGDP 1949
QY 176 OPTMELLCLYPLRLTGGASGCTDFVWAGRVLTCPGDFNLFVLPPTVEOKTRPFTVP 235
DB 1950 GLVPTLVLSVYNNL-INPPGSGTSAIQVT--VETPSEDFFVMI-----1991
QY 236 NIPLKYLNSRIPNPIEGMSLSPDOTQNVQFQNGRCTIDGQPLGTPV--SVSOLCKFRG 293
DB 1992 -----RAPSSKTVDISIP-----AGLLTTPVLTVGVGNDRWNG 2024
QY 294 RITSGORVLNLTDLGSPMAFAAPAGFPDLGSCDMHIEMSKIPNSSTQNNPIVTNSV 353
DB 2025 QIVGLQ-----PVPGF--STCNRRHNL-----NGSTYG-----2051
QY 354 KPNSQQFVPHLSSITLDENVSSGGDYIGTIQWTS-----PSDGGGANTN--F 399
DB 2052 -----WSSPRFGDIGHRRGSASYPGNNAITNVLOF 2080
QY 400 WKIPDYGSLAE-ASQLAPAVPPGFNEVIVFMASIP--GP-----NOS 441
DB 2081 W-YANAGSAIDNPISQVAPDGFDP-----MSFVPPNGPGIPAAAGVWVGFAIWNNS 2130
QY 442 GSPNLVPCLLPOEYITHFISEQAPIQGEAALLHYVDPTNRNLGFEKLYPGCYLTCVFN- 500
DB 2131 GAPNVTVT---QAYELGF-ATGAP-----GNLQPTNTS-----GSQTVAKSI 2169
QY 501 -SSSTGPGQPLDGVVFASWV 521
DB 2170 YAVVTGTAGNPA-GLFVWASGV 2190

RESULT 9
C48562
coat protein - San Miguel sea lion virus (serotype 4)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: C48562
R:Neill, J.D.
Virus Res. 24, 211-222, 1992
A:Title: Nucleotide sequence of the capsid protein gene of San Miguel sea
      lions.
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Accession: C48562
A:Molecule type: genomic RNA
A:Residues: 1-703 <NEI>
A:Cross-references: GB:M87482; NID:g334886; PIDN:AAAL6220.1; PID:g334888
A>Note: sequence extracted from NCBI backbone (NCBIP:113567)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.9%; Score 285; DB 1; Length 703;
Best Local Similarity 24.1%; Pred. No. 3.2e-12;
Matches 116; Conservative 77; Mismatches 194; Indels 94; Gaps 23;
QY 12 ADGATGAGOLVPE---VNTADPIPDPVAGSSTALATAGVNLIDPMI-----INN 59
DB 153 SDG-PSGAETVEEGTGVVQQQAPAP-PTALATLATASTGK-SVQEQWNTFFSYHTSN- 208
QY 60 FVQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGVGNMRVRVVLGNAFT 119
```

```
DB 209 -----WSTVESQKILYSQALNPSINPYLDHIAKLYSTWGGIDVRFVTSVSGGVF 258
QY 120 AKGVICCVPPGFQ-SRTLSIAQATLPPHHIADVRTLDPEVPLEDVRNVLVYHNDTOPT 178
DB 259 GGLAALLVPPGVEPIESVMLQ---YHVLFDARQTEPVIPTIDIRKTLPHSMDTDT 315
QY 179 MRLLCMLYPLRTGASGCTDFVWAGRVLTCPGDFNLFVLP--TVEQKTRP-TVP 235
DB 316 TKL---VINPYENGVENKTCITVETR-----PSADFTFALLKPPGSLIKHGSISPSDLIP 368
QY 236 NIPLKYLNSRIPNPIEGMSLSPDOTQNVQFQNGRCTIDGQPLG--TTPVSVSOLCKFRG 294
DB 369 RNSAHMGN-RWMSTISGFSQVPRVQS-----NRHFDSDSTTGTWSTPYVPIETKIOQK 423
QY 295 ITSQORVLNLTDLGSPMAFAAPAGFPDLGSCDMHIEMSKIPN--SSTQNNPIVTNS 352
DB 424 VGSNNKFWHVIDTD---KALVPGIPDGMFD-----TTIPDETAKATNGNFSYGES 469
QY 353 VKPNSQQFVPHLSSITLDENVSSGGDYI-GTIQWTSPPSDSGGANTNFWKIPDYGSSLA 411
DB 470 YRAGSTTIKPN-----ENSTHFKGTICGTLSTVEIPEN-----DEQIKTE 511
QY 412 ASQLAPAVPPGFNEVIVFMASIPGPNQSGSPNLVPCLLPOEYITHFISEQAPIQGEA 471
DB 512 AEKKSQTMV-----VVTADFKDITVKKQHK-----ISPOKLVVYFDGPEKDLTMSAT 558
QY 472 L 472
DB 559 L 559

RESULT 10
A48562
coat protein - San Miguel sea lion virus (serotype 1)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: A48562
R:Neill, J.D.
Virus Res. 24, 211-222, 1992
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea
      lions.
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Accession: A48562
A:Molecule type: genomic RNA
A:Residues: 1-702 <NEI>
A:Cross-references: GB:M87481; NID:g334882; PIDN:AAAL6217.1; PID:g334884
A>Note: sequence extracted from NCBI backbone (NCBIP:113564, NCBIP:113565)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.7%; Score 281; DB 1; Length 702;
Best Local Similarity 21.8%; Pred. No. 6.1e-12;
Matches 116; Conservative 85; Mismatches 219; Indels 112; Gaps 21;
QY 2 MMASKDAP-----SADGATGAGOLVPEVNTA---DPIDPVPAGSSTALATAGVNL 52
DB 137 LQSTVDTKPDVFRRAESDGGADIVTEEGTVVQQQPV---PAQSALTTLAASGTGKV 193
QY 53 DPWIIINPVQAPQGEFTI-----SPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNG 102
DB 194 D-----CEWTTFFSVHTAVNWSSTEAQSKILFSRALSPELNPYLRRHISLYST 241
QY 103 WGNMRVRVLAGNAFTAGKVIICVPPGFSQRTLSIAQATL--PPHVIADVRTLDPEV 160
DB 242 WSGGIDVRFVTSVSGGVFGGLAALIVPPGIE---PVESPTMLQYPHVLFDARQTEPVI 297
QY 161 PLEDVRNVLVYHNDTOPTMRLLCMLYPLRTGASGCTDFVWAGRVLTCPGDFNLF 220
DB 298 TIPDIRKTLXHSMDTDTTTRLVIMVYNELINPYEOPSEPKS-SCSITVETPSSDFTFSL 356
```

QY 221 VPPTVEQTRPPTVNP- --LKYLNSRIPNPIEGMSLSPDQTNVQFQNGRCTIDGQPL 278
Db 357 KPQSLKXGSPDLSIPNRRHMGNRWMSIDGFFVQPRVFSQ- --NRHFDSDSTTT 412
QY 279 GTT-----PVSVQLCKFRGRTSQRVNLNLTDLGSPFAAPAPAGFPDL----- 326
Db 413 GMSPTYPIEVT-----LEKLDRGQGVFKVTDTEKS-----LVPLGPDGWDPTIPTAMT 463
QY 327 ---GSCDWHIEMSKIPNSSTQNNPIVTSVKPNSQFVPHLSSTITLDENVSS---GGDYI 380
Db 464 ASGNFYDYVAERYTNNCT-----HPKGFYIMGNLTATKVKGSNDL 504
QY 381 GTIQWTSPPSDSGGANTFWKIPDYGSSLAELASQALAPAVYPGFEVIVYFMAVIPGNQ 440
Db 505 GETQOTS-----RTLFSVGVNY-----KDQNTINP--THKITSLSLVYDANNV---S 547
QY 441 SCSPLNVPCLLQOEYTHISQAPIQGEAALLHVVD--PDNRNLGSEFKLY 490
Db 548 AATAKTTTWHSTMSHGLVYLDESVPVGSSTKVVRIATLPEAFTNGGNFPVF 599
coat protein - feline calicivirus (strain Japanese F4)
N;Alternate names: capsid protein
C;Species: feline calicivirus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: B40481
R;Tobya, Y.; Taniguchi, Y.; Takahashi, E.; Utogawa, E.; Takeda, N.; Miyamura, K.; Yamazaki
Virology 183, 810-814, 1991
A;Title: Sequence analysis of the 3'-end of feline calicivirus genome.
A;Reference number: A40481; MUID:91306470; PMID:1853578
A;Molecule type: genomic RNA
A;Residues: 1-668 <TOH>
A;Cross-references: GB:D90357; NID:G221264; PIDN:BAAL4371.1; PID:G221266
C;Superfamily: feline calicivirus coat protein
C;Keywords: capsid protein; coat protein; glycoprotein
F;177,301,304,399,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.5%; Score 274; DB 1; Length 668;
Best Local Similarity 26.5%; Pred. No. 1.8e-11;
Matches 113; Conservative 63; Mismatches 161; Indels 90; Gaps 20;

QY 23 PEVNTA-DPIPIDVAGSSTALATAGQVNLIDPWIIINNFVQAPQGEFTI-----SPNNT 75
Db 133 PEQGTAVGGVTAEPSAQMSAADMASGKSDSEW-----EAFSFTSVNWSSTSET 103
QY 76 PGDVLFDLQGLPHLNPFLSHLSQMYNGVGNMRVVLVAGNAFTAGKVIICCVPPGFQS- 134
Db 184 QCKILFKQSLGFLNPLYLHLSKLYVAVWSGSIEVRSISGSGVFGKLAIVVPPGVDPV 243
QY 135 RTLSIAQATLPHVIADVRTLDPEVPLEDVNRVLYHNNDTQPTWRLLCMLYTPLRGGA 194
Db 244 QSTGMLQ---YPHVLFDAQVPEVIFTPLDRLSTLYHVMSTDTTSLVIMVYNDLINPYA 300
QY 195 SGGTDSFVVAGRVLTCPGPDNELFLVPTVEQKTRPTVNPILKYLNSRIPNPIEGM 254
Db 301 NDSNSSGGCIV--TVETKPGDPFKHLLKPPG-----SVLTGHSIFSDILPK 344
QY 255 SILS-----PDQTONY-----QFO--NGRCTIDGQPLG--TTPVSVSQLCKFRGRTSQGR 300
Db 345 SSSLWIGNRYMTDITDFVIRPVPFQANRHFQNETAGWSTP-----RFR-PIT----- 392
QY 301 VNLNLTDLGSPMAFAAP-----APAGFPDLGSCDWHIEMSKIPNSSTQNNPIVTSVK 354
Db 393 -ITISEKNGSKLIGVATDYIIPGIPDGWPDPTTIAEDLTPAGDYAITSGNGNDITGCS 450
QY 355 PMSQQFVPHLSITLDENVSSGGDYI-GTIQ--WTSPPSDSGGANTFWKIPDYGVSSLA 411
Db 451 -----YDTRAAVVKNTNFRGMVIGCSLQRAW---GDKKISNTAF-----ITTAIRD 493
QY 412 ASQLAPA 418

Db 494 GNEIKES 500
RESULT 12
QY 2354
Db 2354
coat protein - feline calicivirus (strain NADC)
N;Alternate names: capsid protein
C;Species: feline calicivirus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2354
R;Seal, B.S.; Ricipath, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A;Title: Analysis of feline calicivirus capsid protein genes: identification of variable
A;Reference number: JQ2354; MUID:94065683; PMID:7504075
A;Accession: JQ2354
A;Molecule type: mRNA
A;Residues: 1-668 <SEA>
A;Cross-references: GB:L09718; NID:G305104; PIDN:AAAL6485.1; PID:G305105
C;Superfamily: feline calicivirus coat protein
C;Keywords: capsid protein; coat protein

Query Match 9.4%; Score 272.5; DB 2; Length 668;
Best Local Similarity 26.7%; Pred. No. 2.2e-11;
Matches 111; Conservative 59; Mismatches 156; Indels 89; Gaps 20;

QY 34 DPVAGSSTALATAGQVNLIDPWIIINNFVQAPQGEFTI-----SPNNTPGDVLFDLQGLP 87
Db 145 EPSAQMSTAAADMATGKSDSEW-----EAFSFTSVNWSSTSETQCKILFKQSLG 195
QY 88 HLNPFHLSHLSQMYNGVGNMRVVLVAGNAFTAGKVIICCVPPGFQS-RTLSIAQATLFP 146
Db 196 LLNPLYLHLSKLYVAVWSGSIEVRSISGSGVFGKLAIVVPPGVDPVQSTGMLQ---YP 252
QY 147 HVIADVRTLDPEVPLEDVNRVLYHNNDTQPTWRLLCMLYTPLRGTGASGRTSFFVAGR 206
Db 253 HVLFDARQVDVPIFSIPDLRSTLYHLMPTDTTSLVIMVYNDLINPYANDSNSSGGCIV-T 311
QY 207 VLTCPGPDNELFLVPTVEQKTRPTVNPILKYLNSRIPNPIEGMSL-----P 258
Db 312 VETKPGDPFKHLLKPPG-----SMLTGHGSVPDLIPKSSLSLWIGNRYWS 356
QY 259 DQTONV-----QFO--NGRCTIDGQPLG--TTPVSVSQLCKFRGRTSQRVNLNLTDLGSP 312
Db 357 DITDFVVPVFPVQANRHFQNETAGWSTP-----RFR-PIT-----ITISEKSGKL 403
QY 313 -----MAFAAPA-PAGFPDLGSCDWHIEMSKIPNSSTQNNPIVTSVKPNSQFVPHLS 366
Db 404 GIGVATDYVPGIPDGWPDPTTIAEDLTPAGDYAITSGNGNDITGCS-----EYDST 454
QY 367 ITLDENVSSGGDYI-GTIQ--WTSPPSDSGGANTFWKIPDYGVSSLAELASQLAPA 418
Db 455 EVIKNTNFRGMVIGCSLQRAW---GDKKISNTAF-----ITTAIKGNKIRPS 500

RESULT 13
VCWRF9
coat protein - feline calicivirus (strain F9)
N;Alternate names: capsid protein
C;Species: feline calicivirus
C;Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 28-Jul-2000
C;Accession: B43382; C45538; PQ0407; S23702
R;Cartier, M.J.; Milon, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.; Turner, P.C.
Virology 190, 443-448, 1992
A;Title: The complete nucleotide sequence of a feline calicivirus.
A;Reference number: A43382; MUID:92410623; PMID:1529544
A;Accession: B43382
A;Molecule type: genomic RNA
A;Residues: 1-671 <CARI>
A;Cross-references: GB:M86379; NID:G323877; PIDN:AAAY79327.1; PID:G323879
R;Cartier, M.J.; Milon, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.
Arch. Virol. 122, 223-235, 1992
A;Title: Identification and sequence determination of the capsid protein gene of feline
A;Reference number: A45538; MUID:92117861; PMID:1731695

A:Accession: C45538
A:Molecule type: Genomic RNA; protein
A:Residues: 1-671 <CAR2>
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879
A:Experimental source: strain F9
A:Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)
R:Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.
J. Gen. Virol. 73, 2429-2433, 1992
A:Title: The cloning, sequencing and expression of a major antigenic region from the feline calicivirus (FCV) strain F9
A:Reference number: PQ0407; MUID:93019069; PMID:1402818
A:Accession: PQ0407
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 217-266, 'SI', 269-336, 'V', 338-395, 'S', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>
A:Cross-references: PIDN:AA23553.1; PID:g257083
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.1%; Score 264.5; DB 1; Length 671;
Best Local Similarity 33.3%; Pred. No. 8.3e-11;
Matches 73; Conservative 32; Mismatches 89; Indels 25; Gaps 7;

QY 13 DCATGAGQLVPEVNT-ADPDPVAGSSTALATAGQVNLDPWIINNFVQAPQGEFTI- 70
DB 127 DGSITA-----PQGTWVGVIAPSAQMSAADMATGKSDSEW-----EAFPSFH 173

QY 71 -----SPNPTGDVLFDLQGLPHLPFLSHLSQMYNGVGNRVVWLAGNAFTAGKVII 125
DB 174 TSVNMTSETQKILFKQSLGPLLPYLEHLAKLVANSGSIEVRFSISGSGVFGKLA 233

QY 126 CCVPPGFS-RTLSIAQATLPHVIVADVRLDPEVPLEDVRNVLVHNNDTQPTMRLLCM 184
DB 234 IVVPPGDPVQSTMLQ---YHVLFDARQVEPVIFCLPDLASTLYHLMSDTRDTTSLVIM 290

QY 195 LYTPLRTGASGTSFVVGAGVLTCPGDFNFELVPP 223
DB 291 VYNDLINPYANDANSSGCIIV-TVETKPGDPFKFHLKPP 328

RESULT 14
VCWVFF
coat protein - feline calicivirus (strain CFI/68 FIV)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C:Date: 30-Jun-1992 #sequence revision 30-Jun-1992 #text_change 20-Aug-1999
A:Accession: A40507; B40507; T09246
R:Neill, J.D.; Reardon, I.M.; Heinrikson, R.L.
J. Virol. 65, 5440-5447, 1991
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline calicivirus (FCV) strain CFI/68
A:Reference number: A40507; MUID:91374597; PMID:1716692
A:Accession: A40507
A:Molecule type: genomic RNA
A:Residues: 1-668 <NE1>
A:Cross-references: GB:M32819; NID:g323874; PIDN:AAA42925.1; PID:g323875
A:Accession: B40507
A:Molecule type: protein
A:Residues: 373-379,403-419,481-489,560-566 <NE2>
R:Neill, J.D.
submitted to the EMBL Data Library, April 1998
A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.
A:Reference number: Z16626
A:Accession: T09246
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: genomic RNA
A:Residues: 1-668 <NE3>
A:Cross-references: EMBL:U13992; NID:g3056875; PIDN:AAC13993.1; PID:g537256
A:Experimental source: strain CFI/68 FIV
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.1%; Score 263.5; DB 1; Length 668;

Best Local Similarity 25.9%; Pred. No. 9.7e-11;
Matches 105; Conservative 59; Mismatches 162; Indels 79; Gaps 18;

QY 23 PEVNT-ADPDPVAGSSTALATAGQVNLDPWIINNFVQAPQGEFTI-----SPNPT 75
DB 133 PEQGTWVGVIAPSAQMSAADMATGKSDSEW-----EAFPSFHTSVNMTSET 183

QY 76 PGDVLFDLQGLPHLPFLSHLSQMYNGVGNRVVWLAGNAFTAGKVIIICVPPGFS- 134
DB 184 QKILFKQSLGPLLPYLEHLAKLVANSGSIVDRFSISGSGVFGKLAIVVPPGIDPV 243

QY 135 RTLSIAQATLPHVIVADVRLDPEVPLEDVRNVLVHNNDTQPTMRLLCMLVTPRTGA 194
DB 244 QSTSMQ---YHVLFDARQVEPVIFCLPDLASTLYHLMSDTRDTTSLVIMVNDLINPYA 300

QY 195 SGGTDSFVVGAGVLTCPGDFNFELVPPTEQKTRPTVNPINIPKYLNS-----RI 247
DB 301 NDSNNGCIIV-TVETKPGDPFKFHLKPP-----GSMLTGHSIPSLDKSSSLWGNRP 354

QY 248 NPPIEGMSLSP---DQTNVQFN---QRTIDGQPLGTTVPVSVSQLCKFRGRTISGORV 301
DB 355 WSDITDFVIRPFVFOANRHFDFNQETAGWSTPRFPI-TITISVKESAKLGIGVATDVI 413

QY 302 LNLTELDGSPMAAPAPAGFPDLGSCDHLMSKIPNSSTQNNPI-----VTNSVKNS 357
DB 414 PGI-----PDGMPD-----TTIPGELV---PVGDIATNGTNDI 445

QY 358 QQFVPHLSITLDENVSSGGYI-GTIQ--WTSPSPDSGGANTNF 399
DB 446 TTAQYDAATEIRNNTFRGMYICGSLQRAW----GDKKISNTAP 486

RESULT 15
JQ2356
capsid protein - feline calicivirus (strain KCD)
C:Species: feline calicivirus
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable regions
A:Reference number: JQ2354; MUID:94065683; PMID:7504075
A:Accession: JQ2356
A:Molecule type: mRNA
A:Residues: 1-668 <SEA>
A:Cross-references: GB:L09719; NID:g305107; PIDN:AAA16487.1; PID:g305108
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein

Query Match 8.8%; Score 253; DB 2; Length 668;
Best Local Similarity 32.7%; Pred. No. 5.3e-10;
Matches 70; Conservative 32; Mismatches 90; Indels 22; Gaps 7;

QY 19 GOLV-PEVNT-ADPDPVAGSSTALATAGQVNLDPWIINNFVQAPQGEFTI----- 70
DB 128 GSIVTPEQGTWVGVIAPSAQMSAADMATGKSDSEW-----EAFPSFHTSVNM 178

QY 71 SPNPTGDVLFDLQGLPHLPFLSHLSQMYNGVGNRVVWLAGNAFTAGKVIIICVPP 130
DB 179 STSETQKILFKQSLGPLLPYLEHLAKLVANSGSIEVRFSISGSGVFGKLAIVVPP 238

QY 131 GFQS-RTLSIAQATLPHVIVADVRLDPEVPLEDVRNVLVHNNDTQPTMRLLCMLYTPL 189
DB 239 GVDPVQSTSMQ---YHVLFDARQVEPVIFCLPDLASTLYHLMSDTRDTTSLVIMVNDL 295

QY 190 RTGASGTSFVVGAGVLTCPGDFNFELVPP 223
DB 296 INPYANDSNSSGCIIV-TVETKPGSDFRHLKPP 328

Search completed: June 1, 2004, 13:55:22
Job time : 12.4966 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:34:56 ; Search time 6.57181 Seconds
(without alignments)
4310.252 Million cell updates/sec

Title: US-09-926-799-4

Perfect score: 2891

Sequence: 1 MMASKDTPSADGATGAGQ.....YOLKPVGTAGPARGRLGVR 544

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2234.5	77.3	546	1 COAT_SOUV3	Q04542 southampton
2	1156	40.0	539	1 COAT_LORDV	P54635 lordsdale v
3	306	10.6	2344	1 POLN_RHDV	P27410 rabbit hemo
4	285	9.9	703	1 COAT_SMSV4	P36285 san miguel
5	281	9.7	702	1 COAT_SMSV1	P36284 san miguel
6	278	9.6	2208	1 POLN_WANCV	Q69014 manchester
7	274	9.5	668	1 COAT_FCVF4	P27405 feline cali
8	264.5	9.1	671	1 COAT_FCVF9	P27406 feline cali
9	263.5	9.1	668	1 COAT_FCVF6	P27404 feline cali
10	148	5.1	2194	1 POLG_HE701	P32537 h genome po
11	135	4.7	2206	1 POLG_POL3L	P03302 poliovirus
12	134	4.6	2175	1 POLG_BOVEV	P12915 b genome po
13	132.5	4.6	2206	1 POLG_POL32	P06209 poliovirus
14	123	4.3	915	1 A180_RAT	Q05140 rattus norv
15	122	4.2	2205	1 POLG_POL2W	P23069 p genome po
16	122	4.2	2207	1 POLG_POL2L	P06210 p genome po
17	121.5	4.2	2194	1 POLG_EC30B	Q9W78 e genome po
18	119.5	4.1	1233	1 IRS1_MOUSE	P35569 mus musculu
19	119	4.1	2153	1 POLG_HRV16	Q82122 h genome po
20	117.5	4.1	1235	1 IRS1_RAT	P35570 rattus norv
21	115.5	4.0	855	1 POLG_HRV3	Q82081 human rhino
22	115.5	4.0	2179	1 POLG_HRV14	P03303 h genome po
23	115	4.0	2196	1 POLG_EC05N	Q9V1J1 e genome po
24	114	3.9	862	1 POLG_EC16H	Q66790 echovirus 1
25	113.5	3.9	760	1 SIX4_HUMAN	Q9U1U6 homo sapien
26	113.5	3.9	2333	1 POLG_FMDV1	P03306 f genome po
27	113	3.9	1075	1 NFG3_HUMAN	P12968 homo sapien
28	113	3.9	2150	1 POLG_HRV2	P04936 h genome po
29	113	3.9	2206	1 POLG_CXA21	P22055 c genome po
30	112.5	3.9	2193	1 POLG_HE71B	Q66478 h genome po
31	112	3.9	907	1 A180_HUMAN	Q60641 homo sapien
32	112	3.9	2208	1 POLH_POL1M	P03300 p genome po
33	111.5	3.9	896	1 CYRB_MOUSE	P26955 mus musculu

34	111.5	3.9	897	1 CYRB_HUMAN	P32927 homo sapien
35	111	3.8	832	1 POLG_HRV1A	P23008 human rhino
36	111	3.8	2185	1 POLG_SVDVH	P16604 s genome po
37	111	3.8	2185	1 POLG_SVDVU	P13900 s genome po
38	110.5	3.8	769	1 TME8_MOUSE	Q9esn3 mus musculu
39	110	3.8	520	1 LAC1_AGABI	Q12541 agaricus bi
40	110	3.8	521	1 RUN2_HUMAN	Q13950 h runt-rela
41	110	3.8	607	1 RUN2_MOUSE	Q08775 m runt-rela
42	109	3.8	2209	1 POLG_POL1S	P03301 p genome po
43	108.5	3.8	1024	1 SZ6L_HUMAN	Q9byh1 homo sapien
44	108	3.7	2193	1 POLG_CX16T	Q9qf31 c genome po
45	107	3.7	1324	1 SALL1_HUMAN	Q9nsac2 homo sapien

ALIGNMENTS

RESULT 1

ID	COAT_SOUV3	STANDARD;	PRT;	546 AA.
AC	Q04542;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	01-OCT-1994 (Rel. 30, Last annotation update)			
DE	Coat protein (Capeid protein).			
OS	Southampton virus (serotype 3).			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;			
OX	Norovirus.			
NCBI_TaxID=37129;				
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93142023; PubMed=8380940;			
RA	Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;			
RT	"Sequence and genome organization of a human small round-structured			
RT	(Norwalk-like) virus."			
RL	Science 259:516-519(1993).			
CC	-!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			

EMBL; L07418; AAA92984.1; -				
DR	PIR; B37491; B37491.			
DR	InterPro; IPR004005; Calici_coat.			
DR	InterPro; IPR008975; Viral_Cap_coat.			
DR	Pfam; PF00915; Calici_coat; 1.			
KW	Coat protein; Glycoprotein.			
FT	CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT	CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT	P06210 p genome po			
FT	CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT	CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).			
SQ	SEQUENCE 546 AA; 58774 MW; C82B2A85AD4B05EA CRC64;			

Query Match	77.3%;	Score 2234.5;	DB 1;	Length 546;
Best Local Similarity	74.9%;	Pred. No. 2.8e-150;		
Matches 408;	Conservative 58;	Mismatches 78;	Indels 1;	Gaps 1;
QY	1	MMASKDTPSADGATGAGQVPEVNTADPTIDPVAGSSTALATAGQVNLDPVNNF 60		
Db	1	MMASKDTPSADGATGAGQVPEVNTADPTIDPVAGSSTALATAGQVNLDPVNNF 60		
QY	61	VOAPQGETTISPNNTPGDVLFDLQGLPHNLPFLSHLSQMYNGWGMVRVVLVLAQNAFTA 120		
Db	61	VQSPQGETTISPNNTPGDVLFDLQGLPHNLPFLSHLSQMYNGWGMVRVVLVLAQNAFTA 120		
QY	121	GKVIICVPPGFGQSRTLSIAQATLPPHVIADVRLDPVEVPLEDVRNVLVYHNDTQPTMR 180		
Db	121	GKIIIVCCVPPGFGTSSSLTIAQATLPPHVIADVRLDPVEVPLEDVRNVLVYHNDTQPTMR 180		

Matches 257; Conservative 78; Mismatches 172; Indels 62; Gaps 17;

QY 1 MMASKDATPSADGATGAGQLVPEVNTADPDPIDVAGSSTALATAGQWNLIDPWIINF 60
 Db 1 LVCMLYTPURTGGSGNSDSFVAGRVLTARSSDSFLVLPPIEQKTRAFVNIPLQ 240
 QY 241 YLSNRIPIPIEGMSLSLSDQTNQVQFNGRCTIDGOPIGTTPSVSCLCKFRGRTSQOR 300
 Db 241 TILSNRFPFLIQGMLSPDASQVQVQFNGRCLIDQGLLTPATSGQLFRVRGKINQAR 300
 QY 301 VLNLDELGSPMAFAAPAGFPDPLGSCDWHIEMSKIPNSSTQNNPIVTSVKPNSQOF 360
 Db 301 TLNLTEVDGKPMFADSPAPVFPDPPGKCDWHIRLSKTPNNPNSGSDPMRSVSQTNVQGF 360
 QY 361 VPHLSITLIDENVS-SCGYIYIGTQWTSPPSSGGANTFMFKIPYDYGSLSLAEASOLAPAV 419
 Db 361 VPHLGSIQDEVENHPTGDIYGTIEWISQPSPTPGTDINLWEIPDYGSSLSQAANLAPPV 420
 QY 420 YPGFNEVIVYFMASTPGNSQSNLVPCLLPPOEYITHFISEQAPIOGEAALLHYVDPD 479
 Db 421 FPPGCEALVYFVSAPFPGNNRSAPNDVPCLLPQEYITHFVSEQAPTMGDALLHYVDPD 480
 QY 480 TNRNLGEFKLYPGYLTCPVNSSTGPOQLPLDGVFVFAFASWSRFFYQLKPVGTAGPARGR 539
 Db 481 TNRNLGEFKLYPGYLTCPVNGVAGPQQLPLNGVFLFVSWSRFFYQLKPVGTASTARGR 540
 QY 540 LGVRR 544
 Db 541 LGVRR 545

RESULT 2

COAT LORDV STANDARD; PRT; 539 AA.

AC PS4635;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Coat protein (Capsid protein).
 OS Lordsdale virus (Human enteric calicivirus).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus
 OC NCBI_TaxID=82658;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96005060; PubMed=7561776;
 RA Dingle K.E., Lambden P.R., Caul E.O., Clarke I.N.;
 RT "Human enteric Caliciviridae: the complete genome sequence and
 RT expression of virus-like particles from a genetic group II small
 RT round structured virus";
 RL J. Gen. Virol. 76:2349-2355(1995).
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.iesb-sib.ch/announce/>
 CC or send an email to license@iesb-sib.ch).
 CC
 CC EMBL; X86557; CAA60255.1;
 DR InterPro; IPR004005; Calici.coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 KW CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 539 AA; 58775 MW; 5B5C63EF72C5FD21 CRC64;

Query Match 40.0%; Score 1156; DB 1; Length 539;
 Best Local Similarity 45.2%; Pred. No. 3.5e-74;

Matches 257; Conservative 78; Mismatches 172; Indels 62; Gaps 17;

QY 1 MMASKDATPSADGATGAGQLVPEVNTADPDPIDVAGSSTALATAGQWNLIDPWIINF 60
 Db 1 LVCMLYTPURTGGSGNSDSFVAGRVLTARSSDSFLVLPPIEQKTRAFVNIPLQ 240
 QY 241 YLSNRIPIPIEGMSLSLSDQTNQVQFNGRCTIDGOPIGTTPSVSCLCKFRGRTSQOR 300
 Db 241 TILSNRFPFLIQGMLSPDASQVQVQFNGRCLIDQGLLTPATSGQLFRVRGKINQAR 300
 QY 301 VLNLDELGSPMAFAAPAGFPDPLGSCDWHIEMSKIPNSSTQNNPIVTSVKPNSQOF 360
 Db 301 TLNLTEVDGKPMFADSPAPVFPDPPGKCDWHIRLSKTPNNPNSGSDPMRSVSQTNVQGF 360
 QY 361 VPHLSITLIDENVS-SCGYIYIGTQWTSPPSSGGANTFMFKIPYDYGSLSLAEASOLAPAV 419
 Db 361 VPHLGSIQDEVENHPTGDIYGTIEWISQPSPTPGTDINLWEIPDYGSSLSQAANLAPPV 420
 QY 420 YPGFNEVIVYFMASTPGNSQSNLVPCLLPPOEYITHFISEQAPIOGEAALLHYVDPD 479
 Db 421 FPPGCEALVYFVSAPFPGNNRSAPNDVPCLLPQEYITHFVSEQAPTMGDALLHYVDPD 480
 QY 480 TNRNLGEFKLYPGYLTCPVNSSTGPOQLPLDGVFVFAFASWSRFFYQLKPVGTAGPARGR 539
 Db 481 TNRNLGEFKLYPGYLTCPVNGVAGPQQLPLNGVFLFVSWSRFFYQLKPVGTASTARGR 540
 QY 540 LGVRR 544
 Db 541 LGVRR 545

RESULT 3

POLN RHDV STANDARD; PRT; 2344 AA.

ID POLN RHDV
 AC P27410;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Non-structural polyprotein (Contains: RNA-directed RNA polymerase
 DE (EC 2.7.7.48); Thiol protease P3C (EC 3.4.22.-); Helicase (2C like
 DE protein); Coat protein).
 DE Rabbit hemorrhagic disease virus (RHDV).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Lagovirus.
 OC NCBI_TaxID=11976;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91361557; PubMed=1840711;
 RA Meyers G., Wirblich C., Thiel H.-J.;
 RT "Rabbit hemorrhagic disease virus -- molecular cloning and nucleotide
 RT sequencing of a calicivirus genome.";
 RL Virology 184:664-676(1991).
 CC -!- FUNCTION: THE P3C CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE
 CC FOR CLEAVAGE AT CERTAIN O/G SITES IN THE POLYPROTEIN.
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA) (N).
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

316 TKL---VINPENGVENKTSITVETR-----PSADFTFALLKPPGSLIKHGSIPSDLIP 368
236 NTLPKYLSNRPINPIEGMSLSPDOTQNVQFONGRCTIDGQPLG-TTPVSVSQLCKPGR 294
369 RNSAHMGN-RWWSITSGFSVQPRVFQS-----NRHFDFTSTTGTGWTSTYYPIEIKIGK 423
295 IYSGQVRLNLTLDGSPFMAFAAPAGPDLGSCDWHIEMSKIPN--SSTQNNPIVITS 352
424 VGSNNKWFVIDTD---KALVPGIDGHPD-----TTIDETKATNGNFSYGES 469
353 VKPNSQOQFPHLSIITLDENVSSGGDYI-GTIOWTSPPSDGGANTNFWKIPDYGSSLA 411
470 YRAGSTTIKEN-----ENSTHFKGTICGTLSTVEIPEN-----DEQIKTE 511
412 ASOLAPAVTPPGNEVIVFMASIPGPNQSGSNLVPCLLPQBYITHFISEQAPQGEAA 471
512 AEKKSQTMV-----VVTADFQKOTIVKPHK-----ISPQKLVVYFDGPEKDLTMSAT 558
472 L 472
559 L 559

RESULT 5
COAT SMSV1 STANDARD; PRT; 702 AA.
ID COAT SMSV1
AC P36284; 1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Coat protein (Capsid protein).
OS San Miguel sea lion virus (serotype 1) (SMSV 1).
OC viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=36406;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92410750; PubMed=1529644;
RA Neill J.D.;
RT "Nucleotide sequence of the capsid protein gene of two serotypes of
San Miguel sea lion virus: identification of conserved and non-
conserved amino acid sequences among calicivirus capsid proteins.";
RL Virus Res. 24:211-222(1992).
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; M87481; AAA16217.1; --
DR PIR; A48562; A48562.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 493 493 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 702 AA; 77850 MW; E6ESA58523DEED37 CRC64;
Query Match 9.7%; Score 281; DB 1; Length 702;
Best Local Similarity 21.8%; Pred. No. 2,8e-12;
Matches 116; Conservative 95; Mismatches 219; Indels 112; Gaps 21;
2 MMASKDATP-----SADGATGAGQLVPEVNTA---DPIPIDFVAGSSTALATAGQVNL 52
137 LQSTVDTKDFVFRASDGGADIVTEBQGTVVQOPV---PAQSALTTLAAASTGKTV 193

53 DPWILNNFQAQGBETI-----SPNTPGDVLFDLQGLPHLNPFLSHLSQMNG 102
194 D-----CEWTTFFSYHTAVNWSSTEAQKILFSRALSPELNPYLRIHSSLYST 241
103 WUGNMRVRVVLAGNAFTAGKVIICCVPPGFQSRITLSIAQATL--FPHVIADVRTLDPEV 160
242 WSGIDVRFVTVSGSGVFGKLAALIVPPGIE----PVESPTMLQYPHVLFDARQTEPVIF 297
161 PLEBVRNVLVHNNDOTPTWRLICMLYPLRTLGASGSDTSFVVAGRVLTCPGDFNLFPL 220
298 TIPDIRKTLVHSMDDTDTTLLVIMVYNELINPYEQSEPKS-SCSITVETRPSSDFTFSL 356
221 VPTVEQKTRPFTVPNIP--LKYLNSRIPNPIEGMSLSPDOTQNVQFONGRCTIDGQPL 278
357 KPGSLLKHGSI PSDLIPNRSRHHMGNRWWSITDGFVQPRVFQS-----NRHFDFTSTTT 412
279 GTT-----PVSVSQLCFRGRITSGQRVNLNLTLDGSPFMAFAAPAGPDL----- 326
413 GWSTPYIPIEVT-----LEKLDRGQYFKVDTDEKS-----LVPGLDGWDPTTIFAMT 463
327 ---GSCDWHIEMSKIPNSSTQNNPIVITSVKPNSQOQFVPHLSITLDENVSS---GGDYI 380
464 ASNGNYDYITVAEYRITNNGT-----HFKGYIMGNLTTKVKGSDNL 504
381 GTIQWTSPPSDSGGANTNFWKIPDYSSLAESAQAPAVYPPGFNEVIVFMASIPGPNQ 440
505 GETQOTS-----RTLFAVGNY-----KDQNTINP-THKITSNLSLVVDANNV---S 547
441 SGSPNLVCLLPQBYITHFISEQAPQGEAAALHYVD--PDTNRNLGEFKLY 490
548 AATKTTTWHSTHGLGVLDSPGSDSTKVVRIATLPEATNGNGNFPVF 599
POLN MANCV STANDARD; PRT; 2208 AA.
ID POLN MANCV
AC Q69014;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: RNA-directed RNA polymerase
(EC 2.7.7.48); Thiol protease 3C (EC 3.4.22.-); Helicase (2C like
protein); Coat protein].
OS Manchester virus (Human enteric calicivirus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Sapovirus.
OX NCBI_TaxID=82659;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95390791; PubMed=7661689;
RA Liu B.L., Clarke I.N., Caul E.O., Lambden P.R.;
RT "Human enteric caliciviruses have a unique genome structure and are
distinct from the Norwalk-like viruses.";
RT Arch. Virol. 140:1345-1356(1995).
RL -!- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).
CC -!- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR
CC THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC

QY 358 QQVPHLSITLDENVSSGDIY-GHIQ--WTSPPSDGGANTNF 399
 DB 446 TTAQYDAATEIRNTNFRGMVIGSLQRAW-----GDKKISNTAF 486

RESULT 10
 POLG_HE701
 ID POLG_HE701 STANDARD; PRT: 2194 AA.
 AC P32537;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 DE Human enterovirus 70 (strain J670/71) (EV 70).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.
 OC NCBI_TaxID=31915;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91037960; PubMed=2172447;
 RA Ryan M.D., Jenkins O., Hughes P.J., Brown A., Knowles N.J., Booth D., Minor P.D., Almond J.W.;
 RA "The complete nucleotide sequence of enterovirus type 70;
 RT relationships with other members of the picornaviridae.";
 RL J. Gen. Virol. 71:2291-2299 (1990)
 CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
 CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the poliovirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA)(N).
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D00820; BAA18891.1; --
 CC PIR; A36253; GNNYE7.
 CC HSP; P03300; IPOV.
 CC MEROPS; C03.UEB; --
 CC -----
 CC InterPro; IPR003593; AAA ATPase.
 CC InterPro; IPR004004; Calici_pol_hel.
 CC InterPro; IPR009003; Cys_ser_trypsin.
 CC InterPro; IPR000199; Pept_3C_picorn.
 CC InterPro; IPR000081; Peptidase_C3.
 CC InterPro; IPR001138; Pico_P1A.
 CC InterPro; IPR002527; Pico_P2B.
 CC InterPro; IPR001676; Rhv.
 CC InterPro; IPR000605; RNA_helicase.
 CC InterPro; IPR007095; RNA_pol_DS_PS.
 CC InterPro; IPR001205; RNA_pol_P3D.
 CC InterPro; IPR007094; RNA_pol_PSVir.
 CC InterPro; IPR008975; Viral_cap_coat.
 CC Pfam; PF00548; Cys-protease-3C; 1.
 CC Pfam; PF02226; Pico_P1A; 1.
 CC Pfam; PF00947; Pico_P2A; 1.
 CC Pfam; PF01552; Pico_P2B; 1.

DR PF000073; rhv; 3.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PD00918; CALICIVIRUSNS.
 DR ProDom; PD001125; Cys_protease_3C; 1.
 DR ProDom; PD001306; Pico_P2A; 1.
 DR ProDom; PD001274; Pico_P2B; 1.
 DR SMART; SM00382; AAA; 1.
 KW Polyprotein; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
 KW Lipoprotein.
 FT CHAIN 2 69 COAT PROTEIN VP4.
 FT CHAIN 70 319 COAT PROTEIN VP2.
 FT CHAIN 320 561 COAT PROTEIN VP3.
 FT CHAIN 552 871 COAT PROTEIN VP1.
 FT CHAIN 872 1014 CORE PROTEIN P2A.
 FT CHAIN 1015 1113 CORE PROTEIN P2B.
 FT CHAIN 1114 1443 CORE PROTEIN P2C.
 FT CHAIN 1444 1532 CORE PROTEIN P3A.
 FT CHAIN 1533 1554 GENOME-LINKED PROTEIN VP3.
 FT CHAIN 1555 1737 PICORNAIN 3C.
 FT CHAIN 1738 2194 RNA-DIRECTED RNA POLYMERASE P3D.
 FT LIPID 2 N-myristoyl glycine (in host) (By similarity).
 FT ACT_SITE 1701 1701 PROTEASE (POTENTIAL).
 FT ACT_SITE 1715 1715 PROTEASE (POTENTIAL).
 SQ SEQUENCE 2194 AA; 244590 MW; 15DBAE96EE06673C CRC64;

Query Match 5.1%; Score 148; DB 1; Length 2194;
 Best Local Similarity 22.1%; Pred. No. 0.031;
 Matches 91; Conservative 53; Mismatches 163; Indels 104; Gaps 19;

QY 57 INNVQAPQGE---FTISPNNTFGDVLF----DLQL-GPHLPFLSHLSQMYNGWYGNMR 108
 DB 374 INNVNDASGVERLRVQISAQSDMDQLFNIPLDIQLEGLPLRTLLGNISRYTHWSGLE 433
 QY 109 VRVVLGNAFTAGKVIICVPPGFSQRTLSIAQATLPHVIADVRLDPEVPLEDVNV 168
 DB 434 MTFMFCGSMFTTKLIICVTPPGSSPT-DRMQAMLATHVWVDFGLQSSITIIIPISGS 492
 QY 169 LVH--NNDTQPMR---LLCMLYPLRTGGASGSDTSFVAGRV-----LTCPCP 213
 DB 493 HYRMNTAKAINANVGYVTCFMQNL--VAPVGAADQCYIVGMVAAKKDFNLRMRDSP 550
 QY 214 DNFLFLVPP-----TVEQ--KTRPFTVPNIPLKYLNSRIPNPEGMS 255
 DB 551 DIGQSAIPEQAATQIGEVKTVANTVESEIKAEGLVIPSINAVETGATSTPEEAI- 609
 QY 256 LSPDQTNVQFQNG--RCTIDGQPLGTPVSVSQCKPGRITSGQRV-----LNLIT 305
 DB 610 ----QTRTVNNHGTAECLVENF-LGRSALVCMRSFEVKNHSTSTSSIQKNFFIWLNTR 664
 QY 306 EL-----DGSFFMAF-----AAPAGAPDPLGSCDWHIENSKIPINSSTQNPIVNSV 353
 DB 665 ELVQIRRMKMFELTYLRFDTEITIVTLRFLSS-----NVSGFLNLTQAWYVPTGAR 719
 QY 354 KPNSQQFVPHLSITLDENVSSGDIYIGTIOMTSPSDSGGANTNFWKIPD 404
 DB 720 KPSSQD-----SFEWQACNFS-----VFFKIND 743

RESULT 11
 POLG_POL3L
 ID POLG_POL3L STANDARD; PRT: 2206 AA.
 AC P03302; Q84783; Q84784; Q84785; Q84786; Q84787; Q84788; Q84789;
 AC Q84790; Q98592; Q98593; Q98594;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; P3A; Genome-linked protein VP3; Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)].

Wed Jun 2 09:13:29 2004

us-09-926-799-4.rsp

OS Poliovirus type 3 (strains P3/Leon/37 and P3/Leon 12A[1]B).
 OC Viruses; serNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Enterovirus.
 OX NCBI_TaxID=12088;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=P3/Leon/37;
 RX MEDLINE=84170338; PubMed=6324200;
 RA Stanway G., Hughes P.J., Mountford R.C., Reeve P., Minor P.D.,
 RA Schild G.C., Almond J.W.;
 RT "Comparison of the complete nucleotide sequences of the genomes of
 RT the neurovirulent poliovirus P3/Leon/37 and its attenuated Sabin
 RT vaccine derivative P3/Leon 12A[1]B."
 RL Proc. Natl. Acad. Sci. U.S.A. 81:1539-1543 (1984).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=P3/Leon 12A[1]B;
 RX MEDLINE=83292339; PubMed=6310508;
 RA Stanway G., Cann A.J., Hauptmann R., Hughes P.J., Clarke L.D.,
 RA Mountford R.C., Minor P.D., Schild G.C., Almond J.W.;
 RT "The nucleotide sequence of poliovirus type 3 Leon 12 alb: comparison
 RT with poliovirus type 1."
 RL Nucleic Acids Res. 11:5629-5643 (1983).
 RN (3)
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 1-878.
 RX MEDLINE=95120467; PubMed=7820548;
 RA Grant R.A., Hitemath C.N., Filman D.J., Syed R., Andries K.,
 RA Hogle J.M.;
 RT "Structures of poliovirus complexes with anti-viral drugs:
 RT implications for viral stability and drug design."
 RL Curr. Biol. 4:784-797 (1994).
 RN (4)
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1-878.
 RX Hitemath C.N., Grant R.A., Filman D.J., Hogle J.M.;
 RA "Binding of the antiviral drug wif51711 to the Sabin strain of type-3
 RT poliovirus - structural comparison with drug-binding in rhinovirus-
 RT 14."
 RL Acta Crystallogr. D 51:473-489 (1995).
 CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
 CC O/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
 CC -1- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the
 CC poliovirus polypeptide. In other picornavirus reactions Glu may be
 CC substituted for Gln, and Ser or Thr for Gly.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}(N).
 CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -1- MISCELLANEOUS: THE SEQUENCE OF STRAIN SABIN VACCINE P3/LEON/37 IS
 CC SHOWN.
 CC -1- MISCELLANEOUS: THE STRAIN SABIN VACCINE P3/LEON/37 IS THE
 CC PROGENITOR OF THE STRAIN SABIN VACCINE P3/LEON 12A[1]B.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; K01392; AAA46914.1; -;
 CC EMBL; X00925; CAA25444.1; -;
 CC PDB; 1PTV; 03-JUN-95.
 CC PDB; 1PVC; 15-SEP-95.
 CC PDB; 1VBA; 11-JUL-96.
 CC PDB; 1VBB; 11-JUL-96.
 CC PDB; 1VBC; 11-JUL-96.
 CC PDB; 1VBE; 11-JUL-96.
 CC MEROPS; C03.001; -;
 CC MEROPS; C03.020; -;

DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR000199; Peptidase_C3.
 DR InterPro; IPR000081; Peptidase_C3.
 DR InterPro; IPR001338; Pico_P1A.
 DR InterPro; IPR002527; Pico_P2B.
 DR InterPro; IPR001676; Rhv.
 DR InterPro; IPR000605; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00548; Cys-protease_3C; 1.
 DR Pfam; PF02226; Pico_P1A; 1.
 DR Pfam; PF00947; Pico_P2A; 1.
 DR Pfam; PF01552; Pico_P2B; 1.
 DR Pfam; PF00073; Rhv; 3.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICIVIRUSNS.
 DR PRODOM; PD001125; Cys_protease_3C; 1.
 DR PRODOM; PD001306; Pico_P2A; 1.
 DR PRODOM; PD001274; Pico_P2B; 1.
 DR Polyprotein; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
 FT 3D-structure; Lipoprotein.
 FT CHAIN 2 69
 FT CHAIN 70 340
 FT CHAIN 341 578
 FT CHAIN 579 878
 FT CHAIN 879 1027
 FT CHAIN 1028 1124
 FT CHAIN 1125 1453
 FT CHAIN 1454 1540
 FT CHAIN 1541 1562
 FT CHAIN 1563 1745
 FT CHAIN 1746 2206
 FT CHAIN 2 69
 FT ACT SITE 1709 1709
 FT ACT SITE 1723 1723
 FT VARIANT 431 431
 FT VARIANT 864 864
 FT VARIANT 908 908
 FT STRAND 4 8
 FT STRAND 25 29
 FT STRAND 36 38
 FT TURN 50 50
 FT HELIX 51 54
 FT HELIX 57 57
 FT STRAND 63 64
 FT TURN 80 81
 FT TURN 83 87
 FT TURN 88 89
 FT STRAND 90 94
 FT STRAND 101 102
 FT HELIX 103 105
 FT TURN 113 115
 FT STRAND 123 123
 FT HELIX 126 128
 FT TURN 129 129
 FT STRAND 133 134
 FT STRAND 138 141
 FT TURN 142 143
 FT STRAND 147 151
 FT TURN 152 152
 FT HELIX 153 155
 FT TURN 156 157
 FT HELIX 159 160
 FT STRAND 168 169
 FT TURN 185 185
 FT STRAND 187 187
 FT TURN 198 198
 FT STRAND 203 203

```

FT TURN 208 209
FT HELIX 213 216
FT STRAND 219 221
FT STRAND 223 224
FT STRAND 226 226
FT TURN 235 236
FT STRAND 242 242
FT STRAND 245 245
FT HELIX 246 248
FT TURN 249 252
FT HELIX 255 260
FT STRAND 263 267
FT TURN 268 270
FT STRAND 273 278
FT STRAND 287 287
FT TURN 289 291
FT STRAND 292 292
FT STRAND 295 307
FT TURN 308 309
FT STRAND 310 310
FT STRAND 314 331
FT STRAND 337 337
FT TURN 348 351
FT TURN 355 356
FT STRAND 363 363
FT TURN 366 367
FT STRAND 379 380
FT STRAND 382 382
FT TURN 383 383
FT HELIX 384 387
FT TURN 388 388
FT STRAND 391 392
FT STRAND 397 397
FT TURN 399 403
FT HELIX 405 408
FT STRAND 410 413
FT TURN 414 415
FT TURN 418 419
FT STRAND 423 426
FT TURN 429 431
FT TURN 433 437
FT HELIX 439 444
FT TURN 445 446
FT STRAND 447 451

Query Match
Best Local Similarity 4.7%; Score 135; DB 1; Length 2206;
Matches 61; Conservative 30; Mismatches 110; Indels 50; Gaps 10;

QY 22 VPEVNTADPIPDVAGSSTALATAGQVNLIDPWIINNFVQAPQG-----EFTISPNNTPT 76
Db 365 IPEFDVTPDIPDGEVKNMELAE-----IDTWIPLNLESTKNTMDMYRVTLSADL 418
QY 77 GVDLFDLQGLPHNLFSLH-----LSQMTNGVGNMVRVVLAGNAFTAGKVIICVPPG 131
Db 419 SQPILCLSPASDPRLSHTMLGELVNLVYTHWAGSLKFTFLFCGSMATGKILVAYAPG 478
QY 132 FQSRTLSIAQTLFPHVIADVRLDPEVPLEDVRNVLVHNDTQPTMR---LLCNLY-- 186
Db 479 AQPPT-SRKEAMLGTHVMDLGLQSSCTVWVWISNVITYRQT-TQDSFTTEGGYISMFYQT 536
QY 187 ---TPLRTGASGGTDSFVYAGRVLTCPGDFNFVLPVPTVEQKTRPTVTENIPKLKLS 243
Db 537 RIVVPLST-----PKSMGLGFSVAC--NDFSRLRLRDTT-----HIS 572
QY 244 NSRINPIEGM 254
Db 573 QSALPQGIEDL 583

RESULT 12
POLG_BOVEV
ID POLG_BOVEV STANDARD; PRT; 2175 AA.

```

```

AC P12915;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2
DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein
DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
OS Bovine enterovirus (strain VG-5-27) (BEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12065;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88117392; PubMed=2828511;
RA Earle J.A.P., Skuce R.A., Fleming C.S., Hoey E.M., Martin S.J.;
RT "The complete nucleotide sequence of a bovine enterovirus.";
RL J. Gen. Virol. 69:253-263(1988).
RN [2]
RX X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 2-840.
RX MEDLINE=95292108; PubMed=7773791;
RA Smyth M., Tate J., Hoey E.M., Lyons C., Martin S.J., Stuart D.;
RT "Implications for viral uncoating from the structure of bovine
RT enterovirus.";
RL Nat. Struct. Biol. 2:224-231(1995).
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-I-Gly bond in the
CC poliovirus polypeptide. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; D00214; BAA24003.1; ALT_SEQ.
CC PIR; A29824; GNNYBE.
CC PDB; 1BEV; 16-SEP-98.
CC MEROPS; C03.00A; -.
CC MEROPS; C03.00B; -.
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR001199; Pept_3C_picorn.
CC InterPro; IPR000081; Peptidase_C3.
CC InterPro; IPR003138; Pico_P1A.
CC InterPro; IPR002527; Pico_P2B.
CC InterPro; IPR001876; Rnv_P2B.
CC InterPro; IPR006005; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_PSVir.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00548; Cys-protease-3C; 1.
CC Pfam; PF02226; Pico_P1A; 1.
CC Pfam; PF00947; Pico_P2A; 1.
CC Pfam; PF01552; Pico_P2B; 1.
CC Pfam; PF00073; rhv_3.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PD00918; CALICVIRUSNS.
CC ProDom; PD001125; Cys_protease_3C; 1.

```

DR ProDom; PD001306; Pico_P2A; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
KW 3D-structure; Lipoprotein.
FT CHAIN 2 69
FT CHAIN 70 317
FT CHAIN 318 559
FT CHAIN 560 840
FT CHAIN 841 990
FT CHAIN 991 1089
FT CHAIN 1090 1419
FT CHAIN 1420 1508
FT CHAIN 1509 1531
FT CHAIN 1532 1714
FT CHAIN 1715 2175
COAT PROTEIN VP4.
COAT PROTEIN VP2.
COAT PROTEIN VP3.
COAT PROTEIN VP1.
CORE PROTEIN P2A.
CORE PROTEIN P2B.
CORE PROTEIN P2C.
CORE PROTEIN P3A.
CORE PROTEIN P3B.
GENOME-LINKED PROTEIN VPG.
PICORNAIN 3C.
RNA-DEPENDENT RNA POLYMERASE P3D.
RNA-myristoyl glycine (in host).
PROTEASE (POTENTIAL).
PROTEASE (POTENTIAL).
FT ACT SITE 1678 1678
FT ACT SITE 1692 1692
FT STRAND 26 26
FT STRAND 36 38
FT TURN 50 50
FT HELIX 51 54
FT STRAND 57 57
FT STRAND 83 87
FT TURN 88 89
FT STRAND 90 94
FT STRAND 101 102
FT HELIX 103 105
FT TURN 113 115
FT STRAND 123 123
FT HELIX 126 128
FT TURN 129 129
FT STRAND 133 134
FT STRAND 138 139
FT TURN 142 143
FT STRAND 147 151
FT TURN 152 152
FT HELIX 153 155
FT TURN 156 157
FT HELIX 159 167
FT STRAND 168 180
FT TURN 185 186
FT STRAND 187 197
FT TURN 198 198
FT HELIX 209 212
FT STRAND 215 217
FT STRAND 219 220
FT HELIX 224 226
FT TURN 227 230
FT STRAND 234 238
FT HELIX 241 245
FT TURN 246 248
FT STRAND 251 256
FT STRAND 265 265
FT TURN 267 269
FT STRAND 270 270
FT STRAND 273 285
FT STRAND 294 309
FT TURN 325 328
FT TURN 332 333
FT STRAND 340 340
FT TURN 343 344
FT STRAND 357 357
FT STRAND 359 359
FT TURN 360 360
FT HELIX 361 364
FT TURN 365 365
FT STRAND 368 369
FT STRAND 375 375
FT TURN 376 377
FT TURN 380 381
FT HELIX 382 384

FT STRAND 389
FT TURN 394
FT STRAND 397
FT TURN 408
FT HELIX 410 413
FT STRAND 415 420
FT TURN 421 422
FT STRAND 423 427
FT STRAND 430 436
FT TURN 440 441
FT STRAND 443 443
FT STRAND 445 451
FT HELIX 461 464
FT TURN 465 466
FT STRAND 468 473
FT STRAND 479 484
FT STRAND 493 494
FT TURN 499 502
FT HELIX 503 506
FT STRAND 510 515
FT STRAND 520 520
FT TURN 523 524
FT STRAND 528 537
FT TURN 539 540
FT STRAND 542 546
FT TURN 550 551
FT STRAND 574 574
FT STRAND 578 578
FT STRAND 582 582
FT TURN 595 596
FT STRAND 597 598
FT HELIX 600 602
FT STRAND 610 613
FT STRAND 619 619
FT STRAND 624 624
FT HELIX 626 628
FT STRAND 629 629
FT STRAND 630 634
FT STRAND 638 641
FT STRAND 644 644
FT TURN 647 649
FT STRAND 651 655
FT HELIX 662 668
FT TURN 669 670
FT STRAND 671 685
FT STRAND 687 687
FT TURN 689 690
FT STRAND 693 693
FT STRAND 699 705
FT TURN 707 708
FT HELIX 718 721
FT STRAND 727 731
FT TURN 732 733

Query Match 4.6%; Score 134; DB 1; Length 2175;
Best Local Similarity 18.4%; Pred. No. 0.3;
Matches 88; Conservative 57; Mismatches 161; Indels 172; Gaps 17;
Qy 31 IPIDVAGSSTALAT-----AGQVNLDPWLNINNVQAPQGEFTISPNTPG- 77
Db 319 LPTKPGGSGYQFWMTTDEDCSPCILPFOPTLEIFIPGKVNLLLETAQVESILEANNREGV 378
Qy 78 -----DVLFDLQL-----GPHLNPFLSHLSQMYNGWGNMVRVVL 114
Db 379 EGVRYVIVSVQDAL-DAQIYALRLGSGPLSSLLGLTAKHVTQWSGVSVEITCMFT 437
Qy 115 GNAFTAGKVIICVPPGFSRSLSTIAQATLFPHVIADVTLPDVEVPL----- 162
Db 438 GTFWTGKVLAYTPFG-GDMPNREAMLGTHVWDFGLQSSITLVIPIWISASHPRGVS 496
Qy 163 -EDVRNVLHNNDTQPTMRLLCMLYTPLTGGASGDTDFWVAGRVLTCPGPDF----- 215
Db 497 NDDVLNYQY-----AAGHTIYQTNWVIP---PGFPNTAGI 531

QY 216 -----NPLFLVPTVQKTRPFTVPNIPKYLNSRIPNPIEGM-----SUS 257
 Db 532 IMMIAAQPNSFRIDKREDMTQTAILQNDPGKMLKDA-IDKQVAGALVAGTTTSHSVA 590
 QY 258 PPOQTQNVQFN-----GCTIDGQPLGTTTPVS 284
 Db 591 TOSTPALQAAETGATSTARDESMIETRTIVPTHGIIHETSVESFFGSSSLVGMPLIATGTS 650
 QY 285 VS-----QLCKFRGRI-----TSQQRVLNLTDELGSPFMAFAAPA 319
 Db 651 ITNWRIDREFVQLRAKMSFTYMRDVEFTIIATSTGQ---NVTTEQHTTYQWYVPP 707
 QY 320 PAGFP-DLQSCDWHIEMSKIPNSSTONNPIVNSKVPNSQQFVPHLSSTLTENVSSG 376
 Db 708 GAPVFSNQDSFQW-----QSGCNPSVFADTDGPPAQFSVPFMSANAYSTVYDG 756

RESULT 13
 POLG_POL32
 ID POLG_POL32 STANDARD; PRT: 2206 AA.
 AC P06209;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C, P3A; Genome-linked protein VP3; Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 OS Poliovirus type 3 (strain 23127).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Enterovirus.
 OX NCBI_TaxID=12087;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87010550; PubMed=3020156;
 RA Hughes P.J., Evans D.M.A., Minor P.D., Schild G.C., Almond J.W., Stanway G.;
 RT "The nucleotide sequence of a type 3 poliovirus isolated during a recent outbreak of poliomyelitis in Finland.";
 RL J. Gen. Virol. 67:2093-2102(1986).
 CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
 CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln--Gly bond in the poliovirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA}(N).
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X04468; CAA28155.1; -.
 DR PIR: A27245; GNNV27.
 DR HSRP: Q84730; LPVC.
 DR MEROPS: C03.001; -.
 DR MEROPS: C03.020; -.
 DR InterPro: IPR003593; AAA ATPase.
 DR InterPro: IPR004004; Calici_pol_hel.
 DR InterPro: IPR009003; Cys_ser_trypsin.
 DR InterPro: IPR000199; Pept_3C_picorn.
 DR InterPro: IPR000081; Peptidase_C3.
 DR InterPro: IPR003138; Pico_P1A.

DR InterPro: IPR002527; Pico_P2B.
 DR InterPro: IPR001676; Rhv.
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR007095; RNA_pol_DS_PS.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR InterPro: IPR007094; RNA_pol_P5vir.
 DR InterPro: IPR008975; Viral_cap_coat.
 DR Pfam: PF00548; Cys-protease-3C; 1.
 DR Pfam: PF02226; Pico_P1A; 1.
 DR Pfam: PF00947; Pico_P2A; 1.
 DR Pfam: PF01552; Pico_P2B; 1.
 DR Pfam: PF00073; rhv; 3.
 DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam: PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICVIRUSNS.
 DR ProDom; PD001125; Cys_protease_3C; 1.
 DR ProDom; PD001306; Pico_P2A; 1.
 DR ProDom; PD001274; Pico_P2B; 1.
 DR SMART; SM00382; AAA; 1.
 KW Polyprotein; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
 KW Lipoprotein.
 FT CHAIN 2 69 COAT PROTEIN VP4.
 FT CHAIN 70 340 COAT PROTEIN VP2.
 FT CHAIN 341 578 COAT PROTEIN VP3.
 FT CHAIN 579 878 COAT PROTEIN VP1.
 FT CHAIN 879 1027 CORE PROTEIN P2A.
 FT CHAIN 1028 1124 CORE PROTEIN P2B.
 FT CHAIN 1125 1453 CORE PROTEIN P2C.
 FT CHAIN 1454 1540 CORE PROTEIN P3A.
 FT CHAIN 1541 1562 GENOME-LINKED PROTEIN VP3.
 FT CHAIN 1563 1745 PICORNAIN 3C.
 FT CHAIN 1746 2206 N-mristoyl glycine (in host) (By similarity)
 FT LIPID 2 2
 FT ACT_SITE 1709 1709 PROTEASE (POTENTIAL).
 FT ACT_SITE 1723 1723 PROTEASE (POTENTIAL).
 SQ SEQUENCE 2206 AA; 245731 MW; F226AD85403C37BA CRC64;
 Query Match 4.6%; Score 132.5; DB 1; Length 2206;
 Best Local Similarity 24.4%; Pred No. 0.39;
 Matches 63; Conservative 37; Mismatches 121; Indels 37; Gaps 13;
 QY 22 VPEVNTADPIPDPVAGSSTALATAGQVNLDPWIIINNFVQAPGSEFTI---SPNNTPGD 78
 Db 365 IPEPDVTPPIDI---PGEVGNVMEAEIDTMIPLNLENTKRTMDMYRVRLSDSANLSGP 421
 QY 79 VLFDLQLGPHLPFLSH-----LSQVNGVGNMVRVVLGNAFTAGKVIICCVPPGQ 133
 Db 422 ILC-LSLSPAADPRLSHTMLGELVLYNTHWAGSLKFTFLFCGSMMATGKLLVAYAPPGAQ 480
 QY 134 SRTLSIAQATLPFHVIADVRLTDPVEVPLEDVRNVLYHNNDTQPTMR---LLCMLY--- 186
 Db 481 PPT-SRKAMGLTHVWDLGLQSSCTMVVPWISNYYRQT-TQDSFTEGGYISMFYQTRI 538
 QY 187 -TPLRTGASGGTDSFVAVAGRVLTCPGDFNPLFLVPPET-VEQKTRPFTVPNIPKYLNSN 244
 Db 539 VVPLSTPRAMD-----MLGFVSAC--NDFSVRLLRDTHTISQAAMPQGVDDLTLEVAQN 590
 QY 245 S----RIPNPIEGMSLSPD 259
 Db 591 ALALSPLKP---QSNLPD 605

RESULT 14
 A180_RAT
 ID A180_RAT STANDARD; PRT: 915 AA.
 AC Q05140;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Clathrin coat assembly protein AP180 (Clathrin coat associated protein).
 DE AP180) (91 kDa synaptosomal-associated protein).

GN SNAP91.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RC TISSUE=Brain;
RX MEDLINE=931178442; PubMed=8440257;
RA Morris S.A., Schroeder S., Flessmann U., Weber K., Ungewickell E.;
RT "Clathrin assembly protein AP180: primary structure, domain
organization and identification of a clathrin binding site.";
RL EMBO J. 12:667-675(1993).
CC -!- FUNCTION: Adaptins are components of the adaptor complexes which
link clathrin to receptors in coated vesicles. Clathrin-associated
protein complexes are believed to interact with the cytoplasmic
tails of membrane proteins, leading to their selection and
concentration. Binding of AP180 to clathrin triskelia induces
their assembly into 60-70 nm coats.
CC -!- SUBCELLULAR LOCATION: Component of the coat surrounding the
cytoplasmic face of coated vesicles in the plasma membrane.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Names=Long;
CC IsoId=Q05140-1; Sequence=Displayed;
CC Names=Short;
CC IsoId=Q05140-2; Sequence=VSP_000173;
CC -!- DOMAIN: Possesses a three domain structure: the N-terminal 300
residues harbor a clathrin binding site, an acidic middle domain
450 residues, interrupted by an Ala-rich segment, and the C-
terminal domain (166 residues).
CC -!- PTM: Phosphorylated (by similarity).
CC -!- SIMILARITY: Contains 1 epsilon N-terminal homology (ENTH) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; X68877; CAA48748.1; -
CC EMBL; X68878; CAA48749.1; -
CC PIR; S36326; S36326.
CC DR; S36327; S36327.
CC InterPro; IPR001026; ENTH.
CC InterPro; IPR008943; PI_bind_N.
CC Pfam; PF01417; ENTH; 1.
CC SMART; SM00273; ENTH; 1.
CC PROSITE; PS50942; ENTH; 1.
CC Coated pits; Alternative splicing; Phosphorylation.
FT DOMAIN 14 145
FT ENTH.
FT DOMAIN 410 413
FT POLY-THR.
FT DOMAIN 535 539
FT POLY-ALA.
FT DOMAIN 547 550
FT POLY-ALA.
FT DOMAIN 678 683
FT POLY-SER.
FT DOMAIN 723 729
FT POLY-SER.
FT VARSPPLIC 614 632
FT Missing (in isoform Short).
FT /FTId=VSP_000173.
FT SEQUENCE 915 AA; 93518 MW; 32EC1B38EC5DF8C0 CRC64;
Query Match 4.3%; Score 123; DB 1; Length 915;
Best Local Similarity 19.6%; Pred. No. 0.56;
Matches 124; Conservative 64; Mismatches 226; Indels 218; Gaps 27;
4 ASKDATPSADGATGAGOLVPEVNTA-----DIPIDP-----VAGSS 40
363 AAAPVPPTGATGAWGDLGCELSAALSSVSPSEAPISDFAPFAPSPPTTTTTPASASASA 422
41 TALATAG--QVNLIDPWINNFVQAPQGEFTISPNNTPGDVLFDLQGLPHLPFFLSLSQ 98
423 TTAATTAATTEVDLFGDAFAASPGAEAPASEGATAPATPAPVAALDACSGNDFFAPS--- 479

QY 99 MYNGWGNVRVVLGNAGFTAGKVIICCVPPGQSRSLTSLAQATLPPHVIADVRLTDPV 158
DB 480 -----EGSAEAAPELDLPAMKP-----PETSAPVVTPTASTAPPV 514
QY 159 EYPLEEDVRNVLYHNNDTQTMRLLCMLYTPURTGASGCTDSFVVAGRVLTCTC---PGPD 214
DB 515 PA-----TAPSPAPTAVATAATTTTAAAAAATTTATTSAATAATAAPPAALD 560
QY 215 -FNFLFLVPTVEQKTRPTVPNIPLKYLNSRIENPIEGMSLSPDQTONVQFQNG---R 270
DB 561 IFGDLUFDSAPEAAASKPDVAPSIDL--FGTDASSPPRGASPVPESSLTADLLSGSGFH 618
QY 271 CTIDGO--EL-----GTFPVSQCLCKFRGRTSGQRVLNLTGLDGSPPMAFAA-- 317
DB 619 CAEDDRHVPLFTTAVDAFAAPSPASTASPAKAE---SSG-----VIDLFGDAFGSSASET 670
QY 318 -PAP-----AGFPDLGSCDWHIEMKIPNSGTONN----- 346
DB 671 QPAPQAVSSSSASADLLAGF--GS---FMAPSTTPTVTPAQNLLQPNFEAFAGTTPSTS 725
QY 347 -----PIVTNSVKPNSQQFVP-----HLSSTILDENVSSG 376
DB 726 SSSSPDPSGDLMLPTWAPSGQAPVGMVPPSPAMSAKGLGSLDLSLASLVNGLIGIST 785
QY 377 GDYIGTIQWTSPPSD--SGGANTNFWK-----IPDYGSLAEASQALAPAVVPPGF 424
DB 786 TSKKGDGLQWNAKEKLTGGAN---WQPKVTPATWSAGVPPQGTVPPTSS-----VPPG- 835
QY 425 NEVIVYPMASIRPGNOSGSPNLVP-----CLLPQBYITHFISEQAIQGEAALLHYVDP 478
DB 836 -----AGAPSVGQPGAGYGHPAGTGMTWMPQPPQVFMFAPMMRPPFPGAAAV----- 881
QY 479 DTNRNLGFEKLYPGGYLTCPVNSSSTGTPQQLP 510
DB 882 -----PGTQLSPSPPTATQSPKPKPP 901
RESULT 15
POLG POL2W
ID POL2W STANDARD; PRT; 2205 AA.
AC P23069;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (PIA); Coat protein VP2
DE (PIB); Coat protein VP3 (PIC); Coat protein VP1 (PID); Picornain 2A
DE (EC 3.4.22.29) (P2A); Core protein P2B; Core protein P2C; Core protein
DE P3A; Genome-linked protein VPG (P3B); Picornain 3C (EC 3.4.22.28)
DE (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].
OS Poliovirus type 2 (strain W-2).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12085;
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE=90155230; PubMed=2154539;
RA Pevear D.C., Oh C.K., Cunningham L.L., Calenoff M., Jubeit B.;
RT "Localization of genomic regions specific for the attenuated, mouse-
adapted poliovirus type 2 strain W-2.";
RL J. Gen. Virol. 71:43-52(1990).
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
O/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-[Gly bond in the
poliovirus polyprotein. In other picornavirus reactions Glu may be
substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
(RNA)(N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
each of which is composed of one copy each of proteins VP1, VP2,
VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:35:27 ; Search time 31.6725 Seconds

(without alignments)
5419.275 Million cell updates/sec

Title: US-09-926-799-4

Perfect score: 2891

Sequence: 1 MMASKDTPSADGATGAGQ.....YQLKPVGTAGPARGRLGVRR 544

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2891	100.0	544	12 Q9DU46	Q9du46 chiba virus
2	2883	99.7	544	12 Q9QTE7	Q9qte7 chiba virus
3	2868	99.2	544	12 Q9I185	Q9i185 human calic
4	2858	98.9	544	12 Q918C5	Q918c5 norwalk-like
5	2853	98.7	544	12 Q91EJ2	Q91ej2 human calic
6	2852	98.7	544	12 Q918D1	Q918d1 norwalk-like
7	2825	97.7	544	12 Q91V41	Q91v41 human calic
8	2812	97.3	531	12 Q8V7J5	Q8v7j5 norwalk-like
9	2792	96.6	544	12 Q91V36	Q91v36 human calic
10	2361.5	81.7	543	12 Q91V43	Q91v43 human calic
11	2347.5	81.2	543	12 Q8JXJ0	Q8jxj0 norwalk-like
12	2346.5	81.2	543	12 Q918B9	Q918b9 norwalk-like
13	2231.5	77.2	546	12 Q8JTM4	Q8jtm4 norwalk-like
14	2226.5	77.0	546	12 Q91V47	Q91v47 human calic
15	2218.5	76.7	544	12 Q8VA27	Q8va27 human calic
16	2096	72.5	530	12 Q83884	Q83884 norwalk vir

17	2093	72.4	530	12 Q9QT39	Q9qt39 norwalk-like
18	2059.5	71.2	540	12 Q9YS14	Q9ys14 norwalk vir
19	2054	71.0	530	12 Q83876	Q83876 norwalk vir
20	2051	70.9	517	12 Q9JH41	Q9jh41 norwalk-like
21	2036.5	70.4	540	12 Q8BCA5	Q8bca5 human calic
22	1998	69.1	539	12 Q8JV55	Q8jv55 norwalk-like
23	1989	68.8	545	12 Q8VA02	Q8va02 human calic
24	1988	68.8	541	12 Q91V42	Q91v42 human calic
25	1982	68.6	545	12 Q918C8	Q918c8 norwalk-like
26	1982	68.6	545	12 Q91W12	Q91w12 norwalk-like
27	1976.5	68.4	544	12 Q66418	Q66418 desert shie
28	1976	68.4	545	12 Q918C2	Q918c2 norwalk-like
29	1964	67.9	539	12 Q91V48	Q91v48 human calic
30	1960	67.8	545	12 Q91H12	Q91h12 human calic
31	1939	67.1	545	12 Q91V45	Q91v45 human calic
32	1895	65.5	543	12 Q8BCA3	Q8bca3 human calic
33	1328.5	46.0	522	12 Q8B558	Q8b558 bovine ente
34	1319.5	45.6	522	12 Q8B425	Q8b425 norwalk-like
35	1312.5	45.4	522	12 Q8B419	Q8b419 norwalk-like
36	1309.5	45.3	522	12 Q8B422	Q8b422 norwalk-like
37	1306.5	45.2	520	12 Q7TBK7	Q7tbk7 bovine ente
38	1306.5	45.2	522	12 Q8V629	Q8v629 norwalk-like
39	1304.5	45.1	520	12 Q7TBL1	Q7tbl1 bovine ente
40	1301.5	45.0	522	12 Q8V628	Q8v628 norwalk-like
41	1268	43.9	519	12 Q9YQ22	Q9yq22 bovine calic
42	1192.5	41.2	556	12 Q9PY75	Q9py75 human calic
43	1192.5	41.2	556	12 Q91V76	Q91v76 norwalk-like
44	1184.5	41.0	556	12 Q91V79	Q91v79 norwalk-like
45	1180.5	40.8	542	12 Q8JMP1	Q8jmp1 norwalk-like

ALIGNMENTS

RESULT 1

Q9DU46 PRELIMINARY; PRT; 544 AA.
AC Q9DU46;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Capsid protein.
OS Chiba virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=99565;
RN [1]_TaxID=99565;
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Chiba 407/1987/JP;
RA Someya Y., Takeda N., Miyamura T.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Chiba 407/1987/JP;
RX MEDLINE=20569531; PubMed=1118371;
RA Someya Y., Takeda N., Miyamura T.;
RT "Complete nucleotide sequence of the Chiba virus genome and functional expression of the 3C-like protease in Escherichia coli.";
RL Virology 278:490-500(2000).
DR EMBL; AB042808; BAB18267.1; --
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58351 MW; 1DB7A4E81FBB246F CRC64;

Query Match 100.0%; Score 2891; DB 12; Length 544;

Best Local Similarity 100.0%; Pred. No. 5.5e-25;

Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMASKDTPSADGATGAGQGVPEVNTADPIPDPVAGSSTALATAGQVNLDPWIINF 60

DB 1 MMASKDTPSADGATGAGQGVPEVNTADPIPDPVAGSSTALATAGQVNLDPWIINF 60

61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAFTA 120
121 GKVIICCVPPGQSRRTLSIAQATLPHVADRTLDPEVEPLEDRVNLVYHNDTQPTMR 180
121 GKVIICCVPPGQSRRTLSIAQATLPHVADRTLDPEVEPLEDRVNLVYHNDTQPTMR 180
181 LLCMLYTPRLTGASGSDTSFVAGRVLTCPGDFNLFVLPPTVEQKTRPTVPNIPK 240
181 LLCMLYTPRLTGASGSDTSFVAGRVLTCPGDFNLFVLPPTVEQKTRPTVPNIPK 240
241 YLSNSRIPNPIEGMSLSPDQTQNVQFQNGRCHTDGQPLGTTVPVSQKCFRGRITSGOR 300
241 YLSNSRIPNPIEGMSLSPDQTQNVQFQNGRCHTDGQPLGTTVPVSQKCFRGRITSGOR 300
301 VLNLTELDSGPFMAFAAPAGAPGPDIGSCDWHIEMSKI PNSTQNNPIVTNSVKPNSQOF 360
301 VLNLTELDSGPFMAFAAPAGAPGPDIGSCDWHIEMSKI PNSTQNNPIVTNSVKPNSQOF 360
361 VPHLSITLTDENVSSGGDYIGTIQWTSPPSDSGGANTFWKIPDYGSSLAESAQLAPAVY 420
361 VPHLSITLTDENVSSGGDYIGTIQWTSPPSDSGGANTFWKIPDYGSSLAESAQLAPAVY 420
421 PPGNEVIVYFMASIPGPNQSGPNLVPCLLPQEIYTHFISEQAPIQGEAALLHYVDPDT 480
421 PPGNEVIVYFMASIPGPNQSGPNLVPCLLPQEIYTHFISEQAPIQGEAALLHYVDPDT 480
481 NRNLGEFKLYPGGYLTCVPNSSSTGPOQLPDLGVFVFWFASWVSRYQLKPVGTAGPARGRL 540
481 NRNLGEFKLYPGGYLTCVPNSSSTGPOQLPDLGVFVFWFASWVSRYQLKPVGTAGPARGRL 540
541 GVR 544
541 GVR 544

RESULT 3
Q91I85 PRELIMINARY; PRT; 544 AA.
AC Q91I85;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Koblenz/433/2000/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=165252;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/Koblenz 433/2000/DE;
RA Kuenkel U., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF394960; AAK72048.1; -
DR InterPro; IPR004005; Calici coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici coat; 1.
SQ SEQUENCE 544 AA; 58222 MW; 1B39A595B733A6A8 CRC64;

Query Match 99.2%; Score 2868; DB 12; Length 544;
Best Local Similarity 99.1%; Pred. No. 3.9e-223;
Matches 539; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MMASKDATPSADGATGAGQLVPEVNTADPIIDPVAGSSTALATAGQVNLDPWLNINNF 60
DB 1 MMASKDATPSADGATGAGQLVPEVNTADPIIDPVAGSSTALATAGQVNLDPWLNINNF 60
QY 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAFTA 120
QY 121 GKVIICCVPPGQSRRTLSIAQATLPHVADRTLDPEVEPLEDRVNLVYHNDTQPTMR 180

61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAFTA 120
61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAFTA 120
121 GKVIICCVPPGQSRRTLSIAQATLPHVADRTLDPEVEPLEDRVNLVYHNDTQPTMR 180
121 GKVIICCVPPGQSRRTLSIAQATLPHVADRTLDPEVEPLEDRVNLVYHNDTQPTMR 180
181 LLCMLYTPRLTGASGSDTSFVAGRVLTCPGDFNLFVLPPTVEQKTRPTVPNIPK 240
181 LLCMLYTPRLTGASGSDTSFVAGRVLTCPGDFNLFVLPPTVEQKTRPTVPNIPK 240
241 YLSNSRIPNPIEGMSLSPDQTQNVQFQNGRCHTDGQPLGTTVPVSQKCFRGRITSGOR 300
241 YLSNSRIPNPIEGMSLSPDQTQNVQFQNGRCHTDGQPLGTTVPVSQKCFRGRITSGOR 300
301 VLNLTELDSGPFMAFAAPAGAPGPDIGSCDWHIEMSKI PNSTQNNPIVTNSVKPNSQOF 360
301 VLNLTELDSGPFMAFAAPAGAPGPDIGSCDWHIEMSKI PNSTQNNPIVTNSVKPNSQOF 360
361 VPHLSITLTDENVSSGGDYIGTIQWTSPPSDSGGANTFWKIPDYGSSLAESAQLAPAVY 420
361 VPHLSITLTDENVSSGGDYIGTIQWTSPPSDSGGANTFWKIPDYGSSLAESAQLAPAVY 420
421 PPGNEVIVYFMASIPGPNQSGPNLVPCLLPQEIYTHFISEQAPIQGEAALLHYVDPDT 480
421 PPGNEVIVYFMASIPGPNQSGPNLVPCLLPQEIYTHFISEQAPIQGEAALLHYVDPDT 480
481 NRNLGEFKLYPGGYLTCVPNSSSTGPOQLPDLGVFVFWFASWVSRYQLKPVGTAGPARGRL 540
481 NRNLGEFKLYPGGYLTCVPNSSSTGPOQLPDLGVFVFWFASWVSRYQLKPVGTAGPARGRL 540
541 GVR 544
541 GVR 544

RESULT 2
Q9QTE7 PRELIMINARY; PRT; 544 AA.
AC Q9QTE7;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Capsid protein.
OS Chiba virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=99565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Chiba 407;
RA Natori K., Suzuki K., Yamakawa Y., Tatsumi M., Sakae K., Kobayashi S., Shinozaki K., Ishiko H., Miyamura T., Takeda N.;
RT "Expression and self-assembly of capsid proteins of the Chiba virus, a genetically distinct Norwalk-like virus."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022679; BAA82106.1; -
DR InterPro; IPR004005; Calici coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici coat; 1.
SQ SEQUENCE 544 AA; 58323 MW; ABDICIFC4F93D872 CRC64;

Query Match 99.7%; Score 2883; DB 12; Length 544;
Best Local Similarity 99.6%; Pred. No. 2.4e-224;
Matches 542; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MMASKDATPSADGATGAGQLVPEVNTADPIIDPVAGSSTALATAGQVNLDPWLNINNF 60
DB 1 MMASKDATPSADGATGAGQLVPEVNTADPIIDPVAGSSTALATAGQVNLDPWLNINNF 60
QY 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAFTA 120

```
Db 121 GKVIICVPPGFSRSLTAQATLPHVIAVRLDPEVPLEDVRLVYHNDTQPTWR 180
QY 181 LLCMLYPLRTGASGGTDSFVAGRVLTCPGPDFNLFVLPPTVEQKTRPPTVPNIPK 240
Db 181 LLCMLYPLRTGASGGTDSFVAGRVLTCPGPDFNLFVLPPTVEQKTRPPTVPNIPK 240
QY 241 YLSNSRIPNPIEGMSLSPDQTNVQFNGRCITDQGLTTPVSVQSLCKFRGRTISGQR 300
Db 241 YLSNSRIPNPIEGMSLSPDQTNVQFNGRCITDQGLTTPVSVQSLCKFRGRTISGQR 300
QY 301 VLNLTELDGSPFMAFAPAGPDLGSCDWHIEMSKIPNSSTONNPIVNSVKNSQOF 360
Db 301 VLNLTELDGSPFMAFAPAGPDLGSCDWHIEMSKIPNSSTONNPIVNSVKNSQOF 360
QY 361 VPHLSSITLDENVSSGGDIYGTIQTWSPPSDGGANTFNWKIPDYGSSLAESAQLAPAY 420
Db 361 VPHLSSITLDENVSSGGDIYGTIQTWSPPSDGGANTFNWKIPDYGSSLAESAQLAPAY 420
QY 421 PGFNEVIVYFMAIPGPNQSGSPNLVPCLLPOEYITHFISEQAPIQGEAALLHYVDPDT 480
Db 421 PGFNEVIVYFMAIPGPNQSGSPNLVPCLLPOEYITHFISEQAPIQGEAALLHYVDPDT 480
QY 481 NRNLGEFKLYPGGYLTCVPSNSTGPOQLPLDGVFVFASWSRFFYQLKPVGTAGPARGRL 540
Db 481 NRNLGEFKLYPGGYLTCVPSNSTGPOQLPLDGVFVFASWSRFFYQLKPVGTAGPARGRL 540
QY 541 GVR 544
Db 541 GVR 544
```

RESULT 4

```
Q918C5 PRELIMINARY; PRT; 544 AA.
AC Q918C5;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Baltimore/277/1993/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171836;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Baltimore/277/1993/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Baltimore/277/1993/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
Sato Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
characterized small round-structured viruses involved in outbreaks of
nonbacterial acute gastroenteritis in the United States, 1990 to
1995.";
RL J. Med. Virol. 53:372-383(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Baltimore/277/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Baltimore/277/1993/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414404; AAL12965.1;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58208 MW; D5170FA4B958A672 CRC64;
```

Query Match 98.9%; Score 2858; DB 12; Length 544;

Best Local Similarity 98.9%; Pred. No. 2.5e-222;

Matches 538; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MMASKDTPSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGQVNLIDPMIINF 60

Db 1 MMASKDTPSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGVNLIDPMIINF 60

QY 61 VOAPQGEFTISNNTPGDLVLDLQGLPHLNPFLSHLSQMYNGWGNMRVRVVLGNAFTA 120

Db 61 VOAPQGEFTISNNTPGDLVLDLQGLPHLNPFLSHLSQMYNGWGNMRVRVVLGNAFTA 120

QY 121 GKVIICVPPGFSRSLTAQATLPHVIAVRLDPEVPLEDVRLVYHNDTQPTWR 180

Db 121 GKVIICVPPGFSRSLTAQATLPHVIAVRLDPEVPLEDVRLVYHNDTQPTWR 180

QY 181 LLCMLYPLRTGASGGTDSFVAGRVLTCPGPDFNLFVLPPTVEQKTRPPTVPNIPK 240

Db 181 LLCMLYPLRTGASGGTDSFVAGRVLTCPGPDFNLFVLPPTVEQKTRPPTVPNIPK 240

QY 241 YLSNSRIPNPIEGMSLSPDQTNVQFNGRCITDQGLTTPVSVQSLCKFRGRTISGQR 300

Db 241 YLSNSRIPNPIEGMSLSPDQTNVQFNGRCITDQGLTTPVSVQSLCKFRGRTISGQR 300

QY 301 VLNLTELDGSPFMAFAPAGPDLGSCDWHIEMSKIPNSSTONNPIVNSVKNSQOF 360

Db 301 VLNLTELDGSPFMAFAPAGPDLGSCDWHIEMSKIPNSSTONNPIVNSVKNSQOF 360

QY 361 VPHLSSITLDENVSSGGDIYGTIQTWSPPSDGGANTFNWKIPDYGSSLAESAQLAPAY 420

Db 361 VPHLSSITLDENVSSGGDIYGTIQTWSPPSDGGANTFNWKIPDYGSSLAESAQLAPAY 420

QY 421 PGFNEVIVYFMAIPGPNQSGSPNLVPCLLPOEYITHFISEQAPIQGEAALLHYVDPDT 480

Db 421 PGFNEVIVYFMAIPGPNQSGSPNLVPCLLPOEYITHFISEQAPIQGEAALLHYVDPDT 480

QY 481 NRNLGEFKLYPGGYLTCVPSNSTGPOQLPLDGVFVFASWSRFFYQLKPVGTAGPARGRL 540

Db 481 NRNLGEFKLYPGGYLTCVPSNSTGPOQLPLDGVFVFASWSRFFYQLKPVGTAGPARGRL 540

QY 541 GVR 544

Db 541 GVR 544

RESULT 5

```
Q91EJ2 PRELIMINARY; PRT; 544 AA.
AC Q91EJ2;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE Nucleocapsid protein.
OS Human calicivirus Hu/NLV/Queen's Arms/Leeds/92/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=159480;
RN [1]
RP SEQUENCE FROM N.A.
RA Clegg C.S., Chamberlain J., Green J., Brown D.W.G., Lewis D.;
RT "Analysis of diversity in the capsid and ORF3 proteins of some
Norwalk-like viruses circulating in the U.K.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
```



```
OC Norovirus.
OX NCBI_TaxID=122920;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/Valetta/95/Malta;
RX MEDLINE=20404883; PubMed=10949950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL; AJ277616; CAB89097.1; -.
DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF00915; Calici_coat; 1.
FT NON_TER 1
FT TER 544
SQ SEQUENCE 544 AA; 58415 MW; B6741846BD82E6D1 CRC64;

Query Match          97.7%; Score 2825; DB 12; Length 544;
Best Local Similarity 98.0%; Pred. No. 1.2e-219;
Matches 533; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 MMWASKDAPPSADGATGAGQLVPEVNTADPIPIDPVAGSSTALATAGQVNLIDPWIINNF 60
DB 1 MMWASKDAPPSADGATGAGQLVPEVNTADPIPIDPVAGSSTALATAGQVNLIDPWIINNF 60
QY 61 VQAPQGEFTISPNPTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGMNMRVRVVLGNAFTA 120
DB 61 VQAPQGEFTISPNPTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGMNMRVRVVLGNAFTA 120
QY 121 GKVIICCVPPGQSRSTLSIAQATLPHVIADVRTLDPEVPLEDVNRVLYHNNDTQPTMR 180
DB 121 GKVIICCVPPGQSRSTLSIAQATLPHVIADVRTLDPEVPLEDVNRVLYHNNDTQPTMR 180
QY 181 LLCMLYTLRTGASGGTDSFVAVAGRLTFCPGDFNLFVLPVPTVEQKTRPFTVPNIPLK 240
DB 181 LLCMLYTLRTGASGGTDSFVAVAGRLTFCPGDFNLFVLPVPTVEQKTRPFTVPNIPLK 240
QY 241 YLSNRIPIPIEGMSLSPDQTNQVQFQNGRCTIDGQPLGTTTPVSVSQLCKFRGRTSGOR 300
DB 241 YLSNRIPIPIEGMSLSPDQTNQVQFQNGRCTIDGQPLGTTTPVSVSQLCKFRGRTSGOK 300
QY 301 VLNLTELDSPPMAFAAPAGAPFDLGSCHWIEMSKIPNSSTQNNPIVNSVKNSQOF 360
DB 301 VLNLTELDSPPMAFAAPAGAPFDLGSCHWIEMSKIPNSSTQNNPIVNSVKNSQOF 360
QY 361 VPHLSITLDENVSSGGDYIGTIQWTSPPDSGGANTFWKIPDYGSSLAELASQLAPAVY 420
DB 361 VPHLSITLDENVSSGGDYIGTIQWTSPPDSGGANTFWKIPDYGSSLAELASQLAPAVY 420
QY 421 PGFNEVIYVFMASIPGNQSGPNLVPCLLPQEVITHFISEQAPIQGEAALLHYVDPT 480
DB 421 PLVFNVEIYVFMACIRGNQSGPNLVPCLSPQEVITHFISEQAPIQGEAALLHYVDPT 480
QY 481 NRNLGEFKLYPGGYLTCVPNSSTGPOQLPLDGVFVFASWSRFRYQLKPVGTAGARGRL 540
DB 481 NRNLGEFKLYPGGYLTCVPNSSTGPOQLPLDGVFVFASWSRFRYQLKPVGTAGARGRL 540
QY 541 GVRV 544
DB 541 GVRV 544

RESULT 8
Q8V7J5 PRELIMINARY; PRT; 531 AA.
AC Q8V7J5;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein (Fragment).
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
```

```
[1]
RN SEQUENCE FROM N.A.
RP Yoda T., Terano Y., Suzuki Y., Yamazaki K., Oishi I., Kuzuguchi T.,
RA Kawamoto H., Utagawa E., Takino K., Shibata T.;
RT "Characterization of Norwalk virus GI Specific Monoclonal Antibodies
RT Generated against Escherichia coli Expressed Capsid Protein and
RT Reactivity of Two GI Specific Monoclonal Antibodies toward GI
RT Recombinant Fragments.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB063325; BAB83500.1; -.
DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF00915; Calici_coat; 1.
FT NON_TER 1
FT TER 531
SQ SEQUENCE 531 AA; 56747 MW; 9688132FDDCFB88F CRC64;

Query Match          97.3%; Score 2812; DB 12; Length 531;
Best Local Similarity 99.2%; Pred. No. 1.3e-218;
Matches 527; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 ATPSDAGATGAGQLVPEVNTADPIPIDPVAGSSTALATAGQVNLIDPWIINNFVQAPQGE 67
DB 1 ATPSDAGATGAGQLVPEVNTADPIPIDPVAGSSTALATAGQVNLIDPWIINNFVQAPQGE 60
QY 68 FTISPNPTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGMNMRVRVVLGNAFTAGKVIICC 127
DB 61 FTISPNPTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGMNMRVRVVLGNAFTAGKVIICC 120
QY 128 VPPGQSRSTLSIAQATLPHVIADVRTLDPEVPLEDVNRVLYHNNDTQPTMRLLCMLYT 187
DB 121 VPPGQSRSTLSIAQATLPHVIADVRTLDPEVPLEDVNRVLYHNNDTQPTMRLLCMLYT 180
QY 188 PLRTGASGGTDSFVAVAGRLTFCPGDFNLFVLPVPTVEQKTRPFTVPNIPLKYLNSRI 247
DB 181 PLRTGASGGTDSFVAVAGRLTFCPGDFNLFVLPVPTVEQKTRPFTVPNIPLKYLNSRI 240
QY 248 NPPIEGMSLSPDQTNQVQFQNGRCTIDGQPLGTTTPVSVSQLCKFRGRTSGORVNLNLTBL 307
DB 241 NPPIEGMSLSPDQTNQVQFQNGRCTIDGQPLGTTTPVSVSQLCKFRGRTSGOKVNLNLTBL 300
QY 308 DGSPMAFAAPAGAPFDLGSCHWIEMSKIPNSSTQNNPIVNSVKNSQOFVPHLSII 367
DB 301 DGSPMAFAAPAGAPFDLGSCHWIEMSKIPNSSTQNNPIVNSVKNSQOFVPHLSII 360
QY 368 TLDENVSSGGDYIGTIQWTSPPDSGGANTFWKIPDYGSSLAELASQLAPAVYPPGFNEV 427
DB 361 TLDENVSSGGDYIGTIQWTSPPDSGGANTFWKIPDYGSSLAELASQLAPAVYPPGFNEV 420
QY 428 IVYFMASIPGNQSGPNLVPCLLPQEVITHFISEQAPIQGEAALLHYVDPTNRNLGEF 487
DB 421 IVYFMASIPGNQSGPNLVPCLLPQEVITHFISEQAPIQGEAALLHYVDPTNRNLGEF 480
QY 488 KLYPGGYLTCVPNSSTGPOQLPLDGVFVFASWSRFRYQLKPVGTAGARG 538
DB 481 KLYPGGYLTCVPNSSTGPOQLPLDGVFVFASWSRFRYQLKPVGTAGARG 531

RESULT 9
Q9IV36 PRELIMINARY; PRT; 544 AA.
AC Q9IV36;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Thistlehall/90/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=122925;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/Thistlehall/90/UK;
```

RA MEDLINE=20404883; PubMed=10949950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL EMBL; AJ277614; CAB89095.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 543 AA; 58216 MW; 11F97A3989CC540 CRC64;

Query Match 81.7%; Score 2361.5; DB 12; Length 543;
Best Local Similarity 79.6%; Pred. No. 3.3e-182;
Matches 434; Conservative 47; Mismatches 61; Indels 3; Gaps 2;

QY 1 MMASKDAPTSADGATGAGQGVPEVNTADPIPDPVAGSSTALATAGQVNLIDPWIINNF 60
Db 1 MMASKDAPTSADGATGAGQGVPEVNTADPIPDPVAGSSTALATAGQVNLIDPWIINNF 60
QY 61 VQAPQGEFTISPNTPGDVLFDLQGLPHLPFLSHLSQMYNGWGNMVRVVLGNAFTA 120
Db 61 VQAPQGEFTISPNTPGDVLFDLQGLPHLPFLSHLSQMYNGWGNMVRVVLGNAFTA 120
QY 121 GKVIICCVPPGFSQRTLSIAQATLFPFHVADVRLDPEVEPLEDVRNLYHNNDTQPTMR 180
Db 121 GKVIICCVPPGFSQRTLSIAQATLFPFHVADVRLDPEVEPLEDVRNLYHNNDTQPTMR 180
QY 181 LLCMLYTLRTGGASGCTDSFVAGRVLTCPGPDNLFVLPPTVEQKTRPFTVNPILK 240
Db 181 LLCMLYTLRTGGASGCTDSFVAGRVLTCPGPDNLFVLPPTVEQKTRPFTVNPILK 240
QY 241 YLSNSRIPNPIEGMSLSPQTNQVQFNGRCITDGOPLGTTTPVSVOLCKFRGRTTSQGR 300
Db 241 YLSNSRIPNPIEGMSLSPQTNQVQFNGRCITDGOPLGTTTPVSVOLCKFRGRTTSQGR 300
QY 301 VLNLTELDSGPFMAFAAPAGFPDLGSCDWHIEMSKI PNSTQNNPIVTNSVKPNSQQF 360
Db 301 VLNLTELDSGPFMAFAAPAGFPDLGSCDWHIEMSKI PNSTQNNPIVTNSVKPNSQQF 360
QY 361 VPHLSSTILDENVSSGGDIYGTIQTWTSPPSD -SGGANTNFWKIPDYGSSLAESAQLAPAV 419
Db 361 VPHLSSTILDENVSSGGDIYGTIQTWTSPPSD -SGGANTNFWKIPDYGSSLAESAQLAPAV 419
QY 420 YPPGNEVIVFMAISIPGPNOSGPNLVPCLLPOEYITHFTISEQAPIQGEAALLHYVDPT 479
Db 420 YPPGNEVIVFMAISIPGPNOSGPNLVPCLLPOEYITHFTISEQAPIQGEAALLHYVDPT 479
QY 481 NRNLGEPKLYPGGYLTCPVNSSSTGPOQLDGVFVFASWVSRFYQLKPVGTAGPARGRL 540
Db 481 NRNLGEPKLYPGGYLTCPVNSSSTGPOQLDGVFVFASWVSRFYQLKPVGTAGPARGRL 540
QY 541 GVRR 544
Db 541 GVRR 544

RESULT 11
Q8JXJ0 PRELIMINARY; PRT; 543 AA.
ID Q8JXJ0;
AC Q8JXJ0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid.
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SzuGI;
RA Katayama K., Kojima S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE FROM N.A.

RX MEDLINE=20404883; PubMed=10949950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL EMBL; AJ277621; CAB89102.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58676 MW; 4D261DFA42621EDE CRC64;

Query Match 96.6%; Score 2792; DB 12; Length 544;
Best Local Similarity 97.1%; Pred. No. 5.5e-127;
Matches 528; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 1 MMASKDAPTSADGATGAGQGVPEVNTADPIPDPVAGSSTALATAGQVNLIDPWIINNF 60
Db 1 MMASKDAPTSADGATGAGQGVPEVNTADPIPDPVAGSSTALATAGQVNLIDPWIINNF 60
QY 61 VQAPQGEFTISPNTPGDVLFDLQGLPHLPFLSHLSQMYNGWGNMVRVVLGNAFTA 120
Db 61 VQAPQGEFTISPNTPGDVLFDLQGLPHLPFLSHLSQMYNGWGNMVRVVLGNAFTA 120
QY 121 GKVIICCVPPGFSQRTLSIAQATLFPFHVADVRLDPEVEPLEDVRNLYHNNDTQPTMR 180
Db 121 GKVIICCVPPGFSQRTLSIAQATLFPFHVADVRLDPEVEPLEDVRNLYHNNDTQPTMR 180
QY 181 LLCMLYTLRTGGASGCTDSFVAGRVLTCPGPDNLFVLPPTVEQKTRPFTVNPILK 240
Db 181 LLCMLYTLRTGGASGCTDSFVAGRVLTCPGPDNLFVLPPTVEQKTRPFTVNPILK 240
QY 241 YLSNSRIPNPIEGMSLSPQTNQVQFNGRCITDGOPLGTTTPVSVOLCKFRGRTTSQGR 300
Db 241 YLSNSRIPNPIEGMSLSPQTNQVQFNGRCITDGOPLGTTTPVSVOLCKFRGRTTSQGR 300
QY 301 VLNLTELDSGPFMAFAAPAGFPDLGSCDWHIEMSKI PNSTQNNPIVTNSVKPNSQQF 360
Db 301 VLNLTELDSGPFMAFAAPAGFPDLGSCDWHIEMSKI PNSTQNNPIVTNSVKPNSQQF 360
QY 361 VPHLSSTILDENVSSGGDIYGTIQTWTSPPSDSGGANTNFWKIPDYGSSLAESAQLAPAV 420
Db 361 VPHLSSTILDENVSSGGDIYGTIQTWTSPPSDSGGANTNFWKIPDYGSSLAESAQLAPAV 420
QY 421 YPPGNEVIVFMAISIPGPNOSGPNLVPCLLPOEYITHFTISEQAPIQGEAALLHYVDPT 480
Db 421 YPPGNEVIVFMAISIPGPNOSGPNLVPCLLPOEYITHFTISEQAPIQGEAALLHYVDPT 480
QY 481 NRNLGEPKLYPGGYLTCPVNSSSTGPOQLDGVFVFASWVSRFYQLKPVGTAGPARGRL 540
Db 481 NRNLGEPKLYPGGYLTCPVNSSSTGPOQLDGVFVFASWVSRFYQLKPVGTAGPARGRL 540
QY 541 GVRR 544
Db 541 GVRR 544

RESULT 10
Q9IV43 PRELIMINARY; PRT; 543 AA.
ID Q9IV43;
AC Q9IV43;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Musgrove/89/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=122918;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/Musgrove/89/UK;
RX MEDLINE=20404883; PubMed=10949950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,


```
Db 421 FPPGFEALVYFVSAPFGPNNSAPNDVPCLLPQBYITHFVSEQAPTMGDAALLHVVD 480
Qy 480 TNRNLGEFKLYPGGYLTCVPNSSSTGPOQLPLDGVFVFSWVSRYQLKPVGTAGPARGR 539
Db 481 TNRNLGEFKLYPEGFVTCMPN--GTGPOQLPLNGVVFVFSWVSRYQLKPVGTASSARGR 538
Qy 540 LGVRR 544
Db 539 LGVRR 543
Db 541 LGVRR 545

RESULT 14
Q9IV47 PRELIMINARY; PRT; 546 AA.
ID Q9IV47
AC Q9IV47;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
DE Capsid protein.
OS Human calicivirus HU/NLV/Whiterose/96/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OC NCBI_TaxID=122914;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=HU/NLV/Whiterose/96/UK;
RC MEDLINE=20404883; PubMed=10949950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL; AJ277610; CAB89091.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 546 AA; 58737 MW; 501F9D2BE2A77B51 CRC64;

Query Match 77.0%; Score 2226.5; DB 12; Length 546;
Best Local Similarity 74.5%; Pred. No. 2.7e-171; Indels 1; Gaps 1;
Matches 406; Conservative 61; Mismatches 77;

Qy 1 MMASKDAPPSADGATGAGQLVPEVNTADPIPDPVAGSSALATAGQVNLDPWIINNF 60
Db 1 MMASKGAPQADGSSGAGQLVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPWIINNF 60
Qy 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHNPFLLSHLSQMYNGWGNMVRVVLGNAFTA 120
Db 61 VQSPQGEFTISPNNTPGDILFDLQGLPHNPFLLSHLSQMYNGWGNMVRVVLGNAFSA 120
Qy 121 GKVIICCVPPGQSRFTLSIAQATLPPHVIADVRTLDPEVEPLEDVRNLYYHNDTQPTMR 180
Db 121 GKIIIVCCVPPGFTSSSLTIAQATLPPHVIADVRTLEPIEMPLEDVRNLYYHNDTQPTMR 180
Qy 181 LLCMLYTLRTGASGGTDSFVAGRVLTCPGDPNLFVLPPTVEQKTRPFTVNPILK 240
Db 181 LVCMLYTLRTGSGSGNSDSFVAGRVLTAPSSDFSFLFVLPPTIEQKTRPFTVNPILQ 240
Qy 241 YLSNSRIPIEIGMSLSPDQNTQVQFNGRCRTIDGQPLGTTTPSVVSQCKPGRITTSQR 300
Db 241 TLSNSRFPFLIQGMILSPDASQVQFNGRCRLIDQLGTTTPATSGQLFRVRKINQAR 300
Qy 301 VLNLTELDGSPMAFAAPAGFPDGLGSCDWHIEMSKI PNSTQNNPIVTVNSVKNPSQOF 360
Db 301 TLNLTEVDGKPFMAFDSFAPVGFDPFGKCDWHMVRVSKTPNNTSSGDPMRSVSVQTNVQGF 360
Qy 361 VPHLSITLDENVV--SGDYIGTIQWTSPPSDGGANTNFWKIPDYGSSLAESOLAPAV 419
Db 361 VPHLSIQDFEVNHTPGDYIGTIEMISQPSPTPLGTDINLWEIPDYGSSLSQANLAPPV 420
Qy 420 YPPGNEVTVYFMASIPGNQSGSNLVPCLLPQBYITHFVSEQAPIQGEAALLHVVD 479
Db 421 FPPGFEALVYFVSAPFGPNNSAPNDVPCLLPQBYITHFVSEQAPTMGDAALLHVVD 480
```


QY 480 TNRNLGEFKLYPGGYLTCVPSNSSSTGPOQLPLDGVFVFASWVSRRFYQLKPVGTAGPARGR 539
Db 481 TNRNLGEFKLYPGGYLTCVPSNGVAGPOQLPLNGVFLFVSWVSRRFYQLKPVGTASTARGR 540
QY 540 LGVRR 544
Db 541 LGVRR 545
RESULT 15
Q8VA27 PRELIMINARY; PRT; 544 AA.
AC Q8VA27;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/C59/99.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=177069;
RN [1]
RP SEQUENCE FROM N.A.
RA Parkas T., Wilton N., Altaye M., Davies D., Chapman F., Thornton S.A.,
RA Jiang X.;
RT "Homologous vs. heterologous immune responses to Norwalk-like viruses
RT among crew members following acute gastroenteritis outbreaks on two US
RT Navy vessels";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF435807; AAL31551.1; -;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58502 MW; D2003426B9B26E35 CRC64;

Query Match 76.7%; Score 2218.5; DB 12; Length 544;
Best Local Similarity 74.4%; Pred. No. 1.2e-170;
Matches 404; Conservative 60; Mismatches 78; Indels 1; Gaps 1;
QY 3 MASKDTPSADGATGAGQLVPEVTADPIPIPVAGSSTALATAGQVNLIDPWIIINNFVQ 62
Db 1 MASKDAPQADGASGAGQLVPEVTADPLPMEFVAGPTAVATAGQVNMIDPWIVNNFVQ 60
QY 63 APQGEFTISPNTPGDVLFDLQGLPHLPFLSHLSQMYNGWGNMRVYVVLAGNAFTAGK 122
Db 61 SPQGEFTISPNTPGDILFDLQGLPHLPFLSHLSQMYNGWGNMRVILLAGNAFSAGK 120
QY 123 VIICVPPGFGQRTLSIAQATLFPVHVIADVRLDPVEVPLEDVNRVLYHNNDTQPTMRL 182
Db 121 IIVCCVPPGFTSSSLTIAQATLFPVHVIADVRLDPVEVPLEDVNRVLYHNDNQPTMRLV 180
QY 183 CMLYPLRTGGASGGTDSFVAVAGVLTCPGDPNLFLLVPPTVEOKTRPFTVPIPLKYL 242
Db 181 CMLYPLRTGGSGGSDSFVAVAGVLTAPSSDFSLFVPPPTIEOKTRAFVPIPLQL 240
QY 243 SNRIPNPIEGMSLSPDQTNQVQFQNGRCTIDQPLGTTTPSVSVQICKFRGRTSGQRYL 302
Db 241 SNRPFSLIQGLMLSPDASQVQFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQARTL 300
QY 303 NUTELDGSFPMAPAPAPAGFPDLSGDWHIEMSKI PNSSQNNPIVTVNSVKNPNSQQFVP 362
Db 301 NLTEVDGKPFMAFDSAPVGFPPDFGKCDWHMRISKTPNNTSSGDPNRSVSQTNVQGFVP 360
QY 363 HLLSITLDENVIS-SGGDYIGTQWTSPPSDSGGANTFWKI PDYGSSLAESAQLAPAVYP 421
Db 361 HLGSIQDFVFNHPTGDIYGTIEWISQPSPTPGTDINLWEIPDYGSSLSQAANLAPPVFP 420
QY 422 PGFNEVIVFMASIPGPNQSGSNLVPCLLPQEYITHFISEQAPIQGEAALLHYVDPDTN 481
Db 421 PGFGEALVYFVAFPGPNRSAPNDVPCLLPQEYITHFVSEQAPTMGDALLHYVDPDTN 480
QY 482 RNLGEFKLYPGGYLTCVPSNSSSTGPOQLPLDGVFVFASWVSRRFYQLKPVGTAGPARGR 541

Db 481 RNLGEFKLYPGGYLTCVPSNGVAGPOQLPLNGVFLFVSWVSRRFYQLKPVGTASTARGRLG 540
QY 542 VRR 544
Db 541 VRR 543
Search completed: June 1, 2004, 13:53:13
Job time : 32.6725 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 46.3034 Seconds
(without alignments)
3289.030 Million cell updates/sec

Title: US-09-926-799-5
Perfect score: 2886
Sequence: 1 MKQASNDANPSDGTANLVP.....VNQFTYLPAGMGAGRRRAL 539

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2886	100.0	539	AAB49704	Aab49704 Small rou
2	1947.5	67.5	548	AAB49705	Aab49705 Small rou
3	1919	66.5	535	AAB49707	Aab49707 Small rou
4	1907.5	66.1	548	Aau91272	Aau91272 Norwalk v
5	1872.5	64.9	540	AAB49706	Aab49706 Small rou
6	1834.5	63.6	550	AAB49709	Aab49709 Small rou
7	1821.5	63.1	542	AAB49708	Aab49708 Small rou
8	1569	54.4	541	AAB49710	Aab49710 Small rou
9	1161.5	40.2	530	AAR50972	Aar50972 Norwalk v
10	1161.5	40.2	530	ADC72176	Adc72176 Norwalk v
11	1157.5	40.1	545	AAB49700	Aab49700 Small rou
12	1151.5	39.9	544	AAB49703	Aab49703 Small rou
13	1148.5	39.8	530	AAB49701	Aab49701 Small rou
14	1106.5	38.3	530	AAR57091	Aar57091 Small rou
15	1105	38.3	546	AAB49702	Aab49702 Small rou
16	322	11.2	669	AAB67461	Aab67461 Amino aci
17	321	11.1	579	AAB08143	Aab08143 RHDV caps
18	303	10.5	623	AAB47044	Aab47044 Feline ca
19	300.5	10.4	547	AAM50107	Aam50107 Feline ca
20	300.5	10.4	671	AAM50108	Aam50108 Feline ca
21	300	10.4	623	AAB47043	Aab47043 Feline ca
22	298	10.3	622	AAB47045	Aab47045 Feline ca
23	297.5	10.3	668	AAB67462	Aab67462 Amino aci
24	293	10.2	668	AAR10686	Aar10686 Feline ca
25	293	10.2	668	AAB04304	Aab04304 Feline ca

26	197	6.8	40	5	AAU91274	Aau91274 Norwalk v
27	184	6.4	40	5	AAU91273	Aau91273 Norwalk v
28	135	4.7	6310	6	ABU39869	Abu39869 Protein e
29	130.5	4.5	934	1	AAU20016	Aau20016 Sequence
30	126.5	4.4	1147	5	ABB76724	Abb76724 Foot and
31	125.5	4.3	31267	6	ABG74786	Abg74786 Human RGS
32	122	4.2	26926	4	AAU05396	Aau05396 Human tit
33	121	4.2	2406	4	ABB64198	Abb64198 Drosophil
34	117	4.1	998	4	ABU53140	Abu53140 Human tse
35	117	4.1	1024	7	ADC11296	Adc11296 Human nov
36	116	4.0	5701	4	ABB36684	Abb36684 Peptide #
37	116	4.0	5701	4	ABB22021	Abb22021 Protein #
38	116	4.0	5701	4	AAM69843	Aam69843 Human bon
39	116	4.0	5701	4	AAM57448	Aam57448 Human bra
40	116	4.0	5701	4	ABG51536	Abg51536 Human liv
41	116	4.0	5701	4	AAM05328	Aam05328 Peptide #
42	114.5	4.0	2975	6	ABM68759	Abm68759 Photorhab
43	112.5	3.9	710	6	ADA54747	Ada54747 Human pro
44	112	3.9	1344	4	AAG93134	Aag93134 C glutami
45	111.5	3.9	776	2	AAM55105	Aam55105 Streptoco

ALIGNMENTS

RESULT 1
AAB49704
ID AAB49704 standard; protein; 539 AA.

AC AAB49704;

XX
DT 04-APR-2001 (first entry)

XX
DE Small round structured virus protein SEQ ID 5.

XX
KW Small round structured virus; SRSV; food poisoning.

XX
OS Small round structured virus.

XX
PN WO200079280-A1.

XX
PD 28-DEC-2000.

XX
PF 22-JUN-2000; 2000WO-JP004095.

XX
PR 22-JUN-1999; 99JP-00175928.

XX
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

XX
PA (DENK-) DENKA SEIKEN KK.

XX
Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

XX
WPI: 2001-080848/09.

XX
N-PSDB; AAF29145.

XX
Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.

XX
Claim 1; Page 50-52; 84pp; Japanese.

XX
This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

XX
Sequence 539 AA;

XX
Query Match 100.0%; Score 2886; DB 4; Length 539;

XX
Best Local Similarity 100.0%; Pred. No. 1.1e-251;

Matches 539; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MKMASNDANPSDGSSTANLVPEVNEVMALPEVVGAAIAAPVAGQONVIDPWIRNFFVOAP 60
DB	1 MKMASNDANPSDGSSTANLVPEVNEVMALPEVVGAAIAAPVAGQONVIDPWIRNFFVOAP 60
QY	61 GGEFTVSPRNAPGEILWSAPGLDNPYLHSLARMYNGYAGGFVQVILAGNAFTAGKII 120
DB	61 GGEFTVSPRNAPGEILWSAPGLDNPYLHSLARMYNGYAGGFVQVILAGNAFTAGKII 120
QY	121 FFAVPPNPFTEGLSPSQVTMFEHIIIVDVRQLEPVLPLPDVNRNFFHYNQSDSTIKLIA 180
DB	121 FFAVPPNPFTEGLSPSQVTMFEHIIIVDVRQLEPVLPLPDVNRNFFHYNQSDSTIKLIA 180
QY	181 MLYTPLRANNAGDDVFTVSCRVLTRPSDFDFIFLVPPTVESRTKPTFTVPIILTVEMSGNS 240
DB	181 MLYTPLRANNAGDDVFTVSCRVLTRPSDFDFIFLVPPTVESRTKPTFTVPIILTVEMSGNS 240
QY	241 RFPPIPLEKLYTGPFSSAFVQVQNGRCTTGDVLLGTTLQLSAVNICTFRGDVTTHIAGSHDYT 300
DB	241 RFPPIPLEKLYTGPFSSAFVQVQNGRCTTGDVLLGTTLQLSAVNICTFRGDVTTHIAGSHDYT 300
QY	301 MNLASQWNSNYDPTETETIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGSHVFTF 360
DB	301 MNLASQWNSNYDPTETETIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGSHVFTF 360
QY	361 KLGSVQYTTDTNNDFTQNTKFTPVGVIQDGNHNEPQWVLPNYSGRTGHNVLHAPA 420
DB	361 KLGSVQYTTDTNNDFTQNTKFTPVGVIQDGNHNEPQWVLPNYSGRTGHNVLHAPA 420
QY	421 VAPTFGEQLLFFRSTMPGCSGYPNMNLDCLLPQEWVQHFCQBAAPQSDVALLRFVNDP 480
DB	421 VAPTFGEQLLFFRSTMPGCSGYPNMNLDCLLPQEWVQHFCQBAAPQSDVALLRFVNDP 480
QY	481 TGRVLFCKLHKSQYVTVHAHTGPHDLVIPNGYFRFDSWVNFYTLAPMNGAGRRAL 539
DB	481 TGRVLFCKLHKSQYVTVHAHTGPHDLVIPNGYFRFDSWVNFYTLAPMNGAGRRAL 539
RESULT 2	
AAB49705	
ID	AAB49705 standard; protein; 548 AA.
XX	
AC	AAB49705;
XX	
DT	04-APR-2001 (first entry)
XX	
DE	Small round structured virus protein SEQ ID 6.
XX	
KW	Small round structured virus; SRSV; food poisoning.
XX	
OS	Small round structured virus.
XX	
PN	WO200079280-A1.
XX	
PD	28-DEC-2000.
XX	
PF	22-JUN-2000; 2000WO-JP004095.
XX	
PR	22-JUN-1999; 99JP-00175928.
XX	
PA	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
FA	(DENK-) DENKA SEIKEN KK.
XX	
PI	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX	
DR	WPI; 2001-080848/09.
DR	N-PSDB; AAF29146.
XX	
FT	Kit for the detection and typing of small round-structured virus (SRSV)
PT	strains for investigation of food poisoning outbreaks, contains
XX	antibodies.
XX	

PS	Claim 1; Page 52-54; 84pp; Japanese.
XX	
CC	This invention relates to a kit for the detection and typing of small
CC	round structured virus (SRSV) strains. The kit contains antibodies
CC	directed against peptides represented in sequences AAB49700 - AAB49710,
CC	which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC	AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC	used for detecting and typing strains of SRSV in order to prevent the
CC	spread of infection and to examine the epidemiology of outbreaks
XX	
SQ	Sequence 548 AA;
Query Match 67.5%; Score 1947.5; DB 4; Length 548;	
Best Local Similarity 66.3%; Pred. No. 8.2e-167;	
Matches 365; Conservative 73; Mismatches 94; Indels 19; Gaps 4;	
QY	1 MKMASNDANPSDGSSTANLVPEVNEVMALPEVVGAAIAAPVAGQONVIDPWIRNFFVOAP 60
DB	1 MKMASNDANPSDGSSTANLVPEVNEVMALPEVVGAAIAAPVAGQONVIDPWIRNFFVOAP 60
QY	61 GGEFTVSPRNAPGEILWSAPGLDNPYLHSLARMYNGYAGGFVQVILAGNAFTAGKII 120
DB	61 GGEFTVSPRNAPGEILWSAPGLDNPYLHSLARMYNGYAGGFVQVILAGNAFTAGKII 120
QY	121 FFAVPPNPFTEGLSPSQVTMFEHIIIVDVRQLEPVLPLPDVNRNFFHYNQSDSTIKLIA 180
DB	121 FFAVPPNPFTEGLSPSQVTMFEHIIIVDVRQLEPVLPLPDVNRNFFHYNQSDSTIKLIA 180
QY	181 MLYTPLRANNAGDDVFTVSCRVLTRPSDFDFIFLVPPTVESRTKPTFTVPIILTVEMSGNS 240
DB	181 MLYTPLRANNAGDDVFTVSCRVLTRPSDFDFIFLVPPTVESRTKPTFTVPIILTVEMSGNS 240
QY	241 RFPPIPLEKLYTGPFSSAFVQVQNGRCTTGDVLLGTTLQLSAVNICTFRGDVTTHIAG----- 295
DB	241 RFPPIPLEKLYTGPFSSAFVQVQNGRCTTGDVLLGTTLQLSAVNICTFRGDVTTHIAG----- 295
QY	296 -----SHDYTMNLASQWNSNYDPTETIPAPLGTDPFVGKIQGMLTQTTREDGSTR 346
DB	296 -----SHDYTMNLASQWNSNYDPTETIPAPLGTDPFVGKIQGMLTQTTREDGSTR 346
QY	301 ADTATPRLFNYYHWHVQDLNLTGTPYDPAEDIPGLTDFRGKVFQVSGRNL-DSTRA 359
DB	301 ADTATPRLFNYYHWHVQDLNLTGTPYDPAEDIPGLTDFRGKVFQVSGRNL-DSTRA 359
QY	347 HKATVSTGSHVHTPKLGSVQYTTDTNNDFTQNTKFTPVGVIQDGNHNEPQWVLPN 406
DB	347 HKATVSTGSHVHTPKLGSVQYTTDTNNDFTQNTKFTPVGVIQDGNHNEPQWVLPN 406
QY	360 HEAKYDTTAGRTPKLGSLEISTD-SDDFDONQFTKFTPVGI---GYDNEAFQWLSLPD 415
DB	360 HEAKYDTTAGRTPKLGSLEISTD-SDDFDONQFTKFTPVGI---GYDNEAFQWLSLPD 415
QY	407 YSGRTGHNHVLAPAVAPTFGEQLLFFRSTMPGCSGYPNMNLDCLLPQEWVQHFCQBAAP 466
DB	407 YSGRTGHNHVLAPAVAPTFGEQLLFFRSTMPGCSGYPNMNLDCLLPQEWVQHFCQBAAP 466
QY	467 AQSDVALLRFVNDPTGRTVLFCKLHKSQYVTVHAHTGPHDLVIPNGYFRFDSWVNFYTL 526
DB	467 AQSDVALLRFVNDPTGRTVLFCKLHKSQYVTVHAHTGPHDLVIPNGYFRFDSWVNFYTL 526
QY	476 AQTVALLRVYVNDPTGRTVLFCKLHKSQYVTVHAHTGPHDLVIPNGYFRFDSWVNFYTL 535
DB	476 AQTVALLRVYVNDPTGRTVLFCKLHKSQYVTVHAHTGPHDLVIPNGYFRFDSWVNFYTL 535
QY	527 APMGNGAGRRR 537
DB	536 APMGTGNGRRR 546
RESULT 3	
AAB49707	
ID	AAB49707 standard; protein; 535 AA.
XX	
AC	AAB49707;
XX	
DT	04-APR-2001 (first entry)
XX	
DE	Small round structured virus protein SEQ ID 8.
XX	
KW	Small round structured virus; SRSV; food poisoning.
XX	
OS	Small round structured virus.
XX	
PN	WO200079280-A1.
XX	

PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000MO-JF004095.
XX
PR 22-JUN-1999; 99JP-00175928.
XX
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
PI
XX WPI: 2001-080848/09.
XX N-PSDB; AAF29148.
DR
DR
XX
XX
PT Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
XX Claim 1; Page 57-59; 84pp; Japanese.
XX
CC This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX
SQ Sequence 535 AA;

Query Match 66.5%; Score 1919; DB 4; Length 535;
Best Local Similarity 65.4%; Pred. No. 3e-164;
Matches 353; Conservative 80; Mismatches 97; Indels 10; Gaps 5;

QY 1 MKMASNDANPSDGSSTANLVPEVNEVMALEPVVGAIAAPVAGQONVIDPWRNNFVQAP 60
DB 1 MKMASNDRAAPSNDGAGLVPEANETMALEPVAGASIAAPLTGQNNIIDPWRILNFVQAP 60

QY 61 GGEFTVSPRNAPGEILMSAPLGDPLNPLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
DB 61 NGEFTVSPRNSPGEVLLNLELGPENLYLAHLARMYNGYAGGFEVQVLLAGNAFTAGKIV 120

QY 121 FAAPVPNPFTEGLSPQVTFPHIIVDVRLQLEPVLIPLDVRNNFYHYNQSDSTIKLIA 180
DB 121 FAAPVPHPFPLENISPCQITMFPFHVIDVRLQLEPVLIPLDVRNNFYHYNQSDSTIKLIA 180

QY 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFILVPPPTVESRTKPFVPIILTVEMSGNS 240
DB 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDFENLYVPPPTLESKTPFTLPIILITIGELTNS 240

QY 241 RFPPIPLEKLYTGPSSAFVQVQNGRGTCTDGVLLGTTLQLSAVNICITFRGDVTH- IAG-SHD 298
DB 241 RFPVPIDELYTSFNSLVQVQNGRCALDGELOQTTLQTLTAICSRFRGINKVSGENHV 300

QY 299 YTMNLASQNSWYDTEELPAPLGTPDFVKGIOGMLTQTTREDGSTRAHKATVSTGSHVF 358
DB 301 WNMQVNTINGTPTDGTDPVAPLGTPTDFSGKLGVLQSORD-HDNACRSHDAVIATNSAKF 359

QY 359 TPKLGSVQYTTNDNFQGTQNTKFTPVQVIO-DGNHONHQPQWVLPNYSRGTGHVHL 417
DB 360 TPKLGAIQGTWEEDDVHINQPTKFTPVGLFENEGFN-----QWTLNYSGALTINMGL 413

QY 418 APAPVPTFGEOLLFPRSTMPGCSGYPNNMLDCLLPQEWVQHFCEBAPQAQSDVALLRFV 477
DB 414 APVAPTFFGEQILLFPRSHIPLKGVADPVIQCLLPQEWIQLHYESAPSQSDVALIRFT 473

QY 478 NPDTCGRVLECKLHSGVYVTAHTGPHDLVIPPNGYFRFDSWVNOFYTLAPMNGAGRRR 537
DB 474 NPDTCGRVLEAKLHRSGYITVANTGSRPIVVPVANGYFRFDTWVNOFYSLAPMGTGNGRRR 533

RESULT 4
AAU91272

ID AAU91272 standard; protein; 548 AA.
XX
AC AAU91272;
XX
DT 18-JUN-2002 (first entry)
XX
DE Norwalk virus associated polynucleotide #1.
XX
XX Norwalk virus; monoclonal antibody; geno group I; geno group II;
KW immunological detection; food; viral infection.
XX
OS Norwalk virus.
XX
PN JP2002020399-A.
XX
PD 23-JAN-2002.
XX
PF 10-JUL-2000; 2000JP-00208151.
XX
PR 10-JUL-2000; 2000JP-00208151.
XX
PA (OSAP) OSAKA PREFECTURE.
PA (NISE-) NIPPON SEIBUTSU KAGAKU CENT KK.
PA (IATR) IATRON LAB INC.
XX
XX WPI: 2002-287412/33.
XX
XX A monoclonal antibody useful in the immunological detection and diagnosis
PT of Norwalk virus infection.
XX
XX Disclosure; Page 12-13; 24pp; Japanese.
XX
CC The invention describes a monoclonal antibody recognising Norwalk virus,
CC a capsid protein of Norwalk virus, or a common antigen epitope on the
CC capsid protein molecule of geno group I and geno group II. The antibody
CC is useful for immunological detection and quantitative analysis of
CC Norwalk virus in foods and the serum of infected patients. This sequence
CC represents a Norwalk virus associated protein described in the invention
XX
SQ Sequence 548 AA;

Query Match 66.1%; Score 1907.5; DB 5; Length 548;
Best Local Similarity 65.2%; Pred. No. 3.4e-163;
Matches 358; Conservative 75; Mismatches 97; Indels 19; Gaps 4;

QY 1 MKMASNDANPSDGSSTANLVPEVNEVMALEPVVGAIAAPVAGQONVIDPWRNNFVQAP 60
DB 1 MKMASNDRAAPSNDGAGLVPEANETMALEPVAGASIAAPLTGQNNIIDPWRILNFVQAP 60

QY 61 GGEFTVSPRNAPGEILMSAPLGDPLNPLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
DB 61 GGEFTVSPRNSPGEVLLNLELGPENLYLAHLARMYNGYAGGFEVQVLLAGNAFTAGKII 120

QY 121 FAAPVPNPFTEGLSPQVTFPHIIVDVRLQLEPVLIPLDVRNNFYHYNQSDSTIKLIA 180
DB 121 FAAPVPNPFIDNLSAAQITMCPHVIVDVRLQLEPVLIPLDVRNNFYHYNQSDSRLRIA 180

QY 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFILVPPPTVESRTKPFVPIILTVEMSGNS 240
DB 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFDFENLYVPPPTLESKTPFTLPIILITIGELTNS 240

QY 241 RFPPIPLEKLYTGPSSAFVQVQNGRGTCTDGVLLGTTLQLSAVNICITFRGDVTHIAG----- 295
DB 241 RFPVPIESLHTSPENIVVQVQNGRVTLDGELMGTTLQTLPSRICAFRGVLTSTSRASDQ 300

QY 296 -----SHDYTMNLASQNSWYDTEELPAPLGTPDFVKGIOGMLTQTTREDGSTRA 346
DB 301 ADIATPRLFNYWHYVQLDNLNGTYPDAEDIPGGLTPTDFRKGKVGASQ-RNPDSITTRA 359

QY 347 HKATVSTGSHVFTEPKLGSVQYTTNDNFQGTQNTKFTPVQVIOQDNHONHQPQWVLPN 406
DB 360 HEAKVDITTAGRTPKLGSLEISTE-SSDFDQDNQPTFRFTPVGI---GVONEADFQWQLSD 415

CC spread of infection and to examine the epidemiology of outbreaks

XX Sequence 550 AA;

Query Match 63.6%; Score 1834.5; DB 4; Length 550;
 Best Local Similarity 62.5%; Pred. No. 1.4e-156;
 Matches 348; Conservative 85; Mismatches 97; Indels 27; Gaps 7;
 QY 1 MKMASNDANPSGSTANLVPEVNEVMALEPVVGAIAAPVAGQNVQVDPWIRNPFVQAP 60
 DB 1 MKMASNDAAAPSDGASLVPEGINETMPLEPVAGASIAAPVAGQNIIDPWIRNPFVQAP 60
 QY 61 GSEFTVSPNAPGELLWSAPLGNPLYLHARMYNGYAGFEVQVILAGNAFTAGKII 120
 DB 61 QGEFTVSPRNSPGEMLNLLELPELNPYLHLSRMNGYAGGVQVVLGNAFTAGKII 120
 QY 121 FFAVPPNFTPEGLSPQVTFPHIIVDVRLQEPVLPLPDVNRNFFVHNQSDSTIKLIA 180
 DB 121 FFAVPPHPPVENISAAQITMCPHVIVDVRLQEPVLPLPDINRPFHYNQENTPRMLVA 180
 QY 181 MLYTLRANAGDDVFTVSCRVLTRPSDFFIPLVPPTVESRTKPFVPIILTVEMSGNS 240
 DB 181 MLYTLPLRA-NSGEDVFTVSCRVLTRPADDFEFLVPPTVESKTKPFILTLGELSGNS 239
 QY 241 RPIPLEKLYTGPSSAFVVPQNGRCTTGVLLGTTQLSANNICTFRG----- 288
 DB 240 RPPAAIDMLYTDPNESIVVPQNGRCTLDGTTQLVPTQICAFRGTLISQTARAADS 299
 QY 289 -DVTHIAGSHDYTMNLASQNSNYDPTBEIPAPLGTDFVGIQGLTQT---TREDG- 342
 DB 300 TSPQBARNPLHVQVKNLDGTYDPTDIPAVLGAIDPKGVGVGASORDVSGQOQGH 359
 QY 343 -STRAHKAIVTSGSVHFTPKLGSVQVTTDTNNDFOTGQNTKFTPVGVIQDGNHNEPQQ 401
 DB 360 YATRAHEAHDITDPKYAPKLGTILIKSG-SDENTNQPIRTPVGM---GDNW--RQ 412
 QY 402 WVLPNYSGRTGHNHILAPAVATFPGEQLLFRSTMPGCGSPNNMLDCLLPQEWVQHFC 461
 DB 413 WELPDYSGRLTLNMLNAPAVSPFGERILFRSIVPSAGGYGSDYIDCLIPQEWVQHFY 472
 QY 462 QEAAPQAQSDVALLRFVNPDTGRVLFECKLHKSQVTVHTGPHDLVIPNGYFRFDSMN 521
 DB 473 QEAAPQSQSAVALRVVNPDTGRNIFEAKLHREGFLTVANCNGNPIVPPNGYFRFEANGN 532
 QY 522 QFYTLAPMGNGAGRRRA 538
 DB 533 QFYTLAPMGSGQRRRA 549

RESULT 7
 AAB49708
 ID AAB49708 standard; protein; 542 AA.

XX AAB49708;

XX 04-APR-2001 (first entry)

XX Small round structured virus protein SEQ ID 9.

XX Small round structured virus; SRSV; food poisoning.

XX Small round structured virus.

XX WO200079280-A1.

XX 28-DEC-2000.

XX 22-JUN-2000; 2000WO-JP004095.

XX 22-JUN-1999; 99JP-00175928.

XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

PA (DENK-) DENKA SEIKEN KK.

XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 XX WPI; 2001-080848/09.
 DR N-PSDB; AAF29149.
 XX Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX Claim 1; Page 59-61; 84pp; Japanese.
 XX This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX Sequence 542 AA;
 SQ
 Query Match 63.1%; Score 1821.5; DB 4; Length 542;
 Best Local Similarity 63.7%; Pred. No. 2e-155;
 Matches 347; Conservative 74; Mismatches 111; Indels 13; Gaps 6;
 QY 1 MKMASNDANPSGSTANLVPEVNEVMALEPVVGAIAAPVAGQNVQVDPWIRNPFVQAP 60
 DB 1 MKMASNDAAAPSDGASLVPEGINETMPLEPVAGASIAAPVAGQNIIDPWIRNPFVQAP 60
 QY 61 GSEFTVSPNAPGELLWSAPLGNPLYLHARMYNGYAGFEVQVILAGNAFTAGKII 120
 DB 61 QGEFTVSPRNSPGEMLNLLELPELNPYLHLSRMNGYAGGVQVVLGNAFTAGKIL 120
 QY 121 FFAVPPNFTPEGLSPQVTFPHIIVDVRLQEPVLPLPDVNRNFFVHNQSDSTIKLIA 180
 DB 121 FFAVPPHPPVENISAAQITMCPHVIVDVRLQEPVLPLPDINRPFHYNQENTPRMLVA 180
 QY 181 MLYTLRANAGDDVFTVSCRVLTRPSDFFIPLVPPTVESRTKPFVPIILTVEMSGNS 240
 DB 181 MLYTLRNSGDDVFTVSCRVLTRPTDFEFLVPPTVESKTKPFILTLISGLTNS 240
 QY 241 RPIPLEKLYTGPSSAFVVPQNGRCTTGVLLGTTQLSANNICTFRGDVTHIAGSH--D 298
 DB 241 RPIPIEQLYTAPNETNVVQCNGRCTLDGELQGTQLLSSAVCFLOGRTVADNGDNMDQ 300
 QY 299 YTMNLASQNSNYDPTBEIPAPLGTDFVGIQGLTQTTRDGDGSTRAHKAT---VSTGS 355
 DB 301 NLLQLYTPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGIYISTTS 360
 QY 356 VHFTPKLGSVQVTTDTNNDFOTGQNTKFTPVGVIQDGNHNEP-QQWVLPNYSGRTGHN 414
 DB 361 GKFTPKIGSIGLHSITEH-VHPNQSRFTPVGAVD---ENTPQQWVLPHYAGSLALN 415
 QY 415 VHLAPAVAPTFFGQQLLFRSTMPGCGSGYPNNN--LDCLLPQEWVQHFCQAPASQDVA 472
 DB 416 TNLAPAVAPTFFGQQLLFRSRVPCVQGLQGDQAFIDCLLPQEWVNHFYQEAAPSQDVA 475
 QY 473 LLRFVNPDTGRVLFECKLHKSQVTVHTGPHDLVIPNGYFRFDSWNVQFVTLAPMGNG 532
 DB 476 LIRVNPDTGRVTLFEAKLHRSQFTVSHGTGAYPLVPPNPNHFRFDSWNVQFVTLAPMGNG 535
 QY 533 AGRRR 537
 DB 536 NRRRR 540
 RESULT 8
 AAB49710
 ID AAB49710 standard; protein; 541 AA.
 XX AAB49710;
 AC AAB49710;
 XX

PR	22-JUN-1999;	99JP-00175928.	
XX	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.		
PA	(DENK-) DENKA SEIKEN KK.		
XX			
PI	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;		
XX			
XX	WPI; 2001-080848/09.		
DR	N-PSDB; AAF29141.		
XX			
XX	Kit for the detection and typing of small round-structured virus (SRSV)		
PT	strains for investigation of food poisoning outbreaks, contains		
PT	antibodies.		
XX			
PS	Claim 1; Page 40-42; 84pp; Japanese.		
XX			
CC	This invention relates to a kit for the detection and typing of small		
CC	round structured virus (SRSV) strains. The kit contains antibodies		
CC	directed against peptides represented in sequences AAB49700 - AAB49710,		
CC	which are each SRSV strain specific. Polynucleotide sequences AAF20141 -		
CC	AAF20151 represent cDNA encoding the strain specific proteins. The kit is		
CC	used for detecting and typing strains of SRSV in order to prevent the		
CC	spread of infection and to examine the epidemiology of outbreaks		
XX			
SQ	Sequence 545 AA;		
	Query Match 40.1%; Score 1157.5; DB 4; Length 545;		
	Best Local Similarity 45.6%; Pred. No. 2.3e-95;		
	Matches 253; Conservative 83; Mismatches 192; Indels 27; Gaps 14;		
Qy	1 MKMASNDANPS-DGST--ANLVPVNN-EVMALPEVVGAIAAPVAGQONVDPWIRNMF 56		
Db	1 MMASKDAPTMDDTSGAGLVPEANTAEPIEPVAGATAAATAGQVNMIDPIMNNY 60		
Qy	57 VQAPGGEFTVSPNAPGEILWSAPGLDLPYLHLARMYNGYAGGFEVQVILAGNAFTA 116		
Db	61 VQAPGGEFTISPNNTPGDIILFDLQGLPHLNFPLSHLAQMYNGVGNMKVLLAGNAFTA 120		
Qy	117 GKIIFAAVPPNPPTGELSPQVTMPHHIIVDVQLEPVLIPLDVPRNNFHYNQSDSTI 176		
Db	121 GKIIISCIPGFAAQNISIAQATMFPHVADRVLEPIEVLPELDRVNLPH-NNDNAPTM 179		
Qy	177 KLIALMYTLPLRA--NNAGDDVFTVSCRVLTPSPDFIFLVPPTVESRTKPTVPILTV 234		
Db	180 RLVCMLYTLPLRASGSSGTDPFVIAGRVLTCTPSPDFSLFVLPNVEQTKPESVNLPL 239		
Qy	235 EEMNSRFPPIPLEKLYTGPSSAFVVQONGRCTTDDGVLGTLQLSAVNICTFRGDVTHIA 294		
Db	240 NTLNSRVPSLISKMWYSRDHQMVQFQNGRVTLDDGQLQGTTPTSASQLCKIRGSVFHAN 299		
Qy	295 GSHDYTMNLASQWNSNYDPTTEIPAPLGTDPFVGKIQQML--TOTTRDGSTRAHKATVS 352		
Db	300 GGNGY--NLTELDGSPYHAFES-PAPIGFPD-LGECODHWEASPTTQFGDVIKQINVK 355		
Qy	353 TGSVHFTPKLSGVQVTTDNNDFOTGQNTKFTPVGVTDGNNHQNPEQQWVLPNYSRTG 412		
Db	356 QESA-FAPHLGTIQADGSDVSNTNMTAKLGWSPVSDG--HRGDVDPWIPRYGSTLT 412		
Qy	413 HNVHLAPAVATPFGEOQLLFFRSTMPGCGSVNNMLDCLLPQEWVQHFQCBAAQSDVA 472		
Db	413 EAAQLAPPIYPGFGCEAIVFFMSDFPIAHGTNGLSVPTCTIPQEFVTHFVNEQAETRGEAA 472		
Qy	473 LLRFVNPDGTVLFECKLHKSQYTV---AHTGPHDVIIPNGVFRFDSWVNOFYTLAP 528		
Db	473 LLHYLDPDTHRLGFBKLYPEGFMTCVNSGTGQTL--PINGVVFVSVNSRYQLKP 530		
Qy	529 MNGA-----GRRRA 538		
Db	531 VGTAGPACRLGIIRS 545		
	RESULT 12		
	AAB49703		
ID	AAB49703 standard; protein; 544 AA.		
XX	AAB49703;		
AC			
XX	04-APR-2001 (first entry)		
DT			
XX	Small round structured virus protein SEQ ID 4.		
DE			
XX	Small round structured virus; SRSV; food poisoning.		
KW			
XX	Small round structured virus.		
OS			
XX	WO200079280-A1.		
BN			
XX	28-DEC-2000.		
PD			
XX	22-JUN-2000; 2000WO-JP004095.		
XX	22-JUN-1999; 99JP-00175928.		
PR	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.		
XX	(DENK-) DENKA SEIKEN KK.		
FA	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;		
PI			
XX	WPI; 2001-080848/09.		
DR	N-PSDB; AAF29144.		
XX			
XX	Kit for the detection and typing of small round-structured virus (SRSV)		
PT	strains for investigation of food poisoning outbreaks, contains		
PT	antibodies.		
XX			
PS	Claim 1; Page 47-49; 84pp; Japanese.		
XX			
CC	This invention relates to a kit for the detection and typing of small		
CC	round structured virus (SRSV) strains. The kit contains antibodies		
CC	directed against peptides represented in sequences AAB49700 - AAB49710,		
CC	which are each SRSV strain specific. Polynucleotide sequences AAF20141 -		
CC	AAF20151 represent cDNA encoding the strain specific proteins. The kit is		
CC	used for detecting and typing strains of SRSV in order to prevent the		
CC	spread of infection and to examine the epidemiology of outbreaks		
XX			
SQ	Sequence 544 AA;		
	Query Match 39.9%; Score 1151.5; DB 4; Length 544;		
	Best Local Similarity 44.6%; Pred. No. 7.9e-95;		
	Matches 254; Conservative 80; Mismatches 175; Indels 61; Gaps 16;		
Qy	1 MKMASNDANPS-DGST--ANLVPVNN-EVMALPEVVGAIAAPVAGQONVDPWIRNMF 56		
Db	1 MMASKDAPTSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGQVNLDPWIINF 60		
Qy	57 VQAPGGEFTVSPNAPGEILWSAPGLDLPYLHLARMYNGYAGGFEVQVILAGNAFTA 116		
Db	61 VQAPGGEFTISPNNTPGDIILFDLQGLPHLNFPLSHLSQMYNGVGNMRVRVVLAGNAFTA 120		
Qy	117 GKIIFAAVPPNPPTGELSPQVTMPHHIIVDVQLEPVLIPLDVPRNNFHYNQSDSTI 176		
Db	121 GKVIICCVPPGQSRSLTSLAQATLPPHVADRVTLDPVEPLEDRVNLVYH-NNDTQPTM 179		
Qy	177 KLIALMYTLPLRANNA--GDDVFTVSCRVLTPSPDFIFLVPPTVESRTKPTVPILTV 234		
Db	180 RLICMLYTLPLRTGGASGGTDSFVAVAGRVLTCTPDPDFNLFLVPPTVEQTKRPTVPNPL 239		
Qy	235 EEMNSRFPPIPLEKLYTGPSSAFVVQONGRCTTDDGVLGTLQLSAVNICTFRGDVTHIA 294		
Db	240 KYLSNSRIENPIEGMSLSPDQNTQVQFQNGRCTIDGQLGTPFVSVSQLCRKRIT--- 296		
Qy	295 GSHDYTMNLASQWNSNYDPTTEIPAPLGTDPFVGKIQQMLTOTTRDGSTRAHKA---TV 351		
Db	297 -SGQRVNLNLTLDGSPF-MAPAAPAPAGFPD-LGSCDWHIEMSKIPNSSTQNPNIVTNSV 353		
Qy	352 STGSVHFTPKL-----GSVQVTTDNNDFOTGQNTKFTPVGVTDGNNH 395		

Db 354 KNSQOFVPHLSITLDENVSSGSDYIGTIQMTSPSD--SGGANTNF----- 399
 Qy 396 QNEPQWVLPNTSGRTGHNHVLAPAVAPTFPGEQLLFFRSTMPG--CSGYPMNMLDCLLP 453
 Db 400 -----WKIPDYGSSILAEASQAPAVYPPGFNEVIVYFWASIPGPNQSGSNL--VPCLLP 452
 Qy 454 QBWVHFQCEAPASQVALLRFVNPDTGRVLFECKLHKSGYVTV-----AHTGPHDLVTP 509
 Db 453 QBYITHFISEQAPIQCEAALLHYDPDTRNLGEFKLPGGYLTCVPNSSSTGPOQL--P 510
 Qy 510 PNGYFRFDSWVNOFYTLAPMGNGAGRRRAL 539
 Db 511 LDGVVFASWSRFFYQLKPVGTAGPARGRL 540

RESULT 13
 AAB49701
 ID AAB49701 standard; protein; 530 AA.
 XX AC AAB49701;
 XX 04-APR-2001 (first entry)
 DT Small round structured virus protein SEQ ID 2.
 DE Small round structured virus; SRSV; food poisoning.
 XX KW Small round structured virus; SRSV; food poisoning.
 XX OS Small round structured virus.
 XX WO200079280-A1.
 PN 28-DEC-2000.
 PD 22-JUN-2000; 2000WO-JP004095.
 PF 22-JUN-1999; 95JP-00175928.
 PR (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.
 XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 XX WPI; 2001-080848/09.
 DR N-PSDB; AAF29142.
 PT Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX Claim 1; Page 42-45; 84pp; Japanese.
 CC This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX Sequence 530 AA;
 SQ

Query Match 39.8%; Score 1148.5; DB 4; Length 530;
 Best Local Similarity 44.8%; Pred. No. 1.4e-94;
 Matches 247; Conservative 85; Mismatches 182; Indels 37; Gaps 14;

Qy 1 MKWASNDANPS--DGST--ANLYPEVN-NEVMALFVVGAAIAAPVAGQONVIDPWIRNF 56
 Db 1 MWASKDATSSVDGASGQLYPEVNASDPLAWDPVAGSSTAVATAGVNPDPWIINF 60
 Qy 57 VQAPGGEFTVSPRNAPGEILWAPLGPDLNPNYLHARMYNGYAGGFEVQVILAGNAFTA 116
 Db 61 VQAPQGEFTISPNNTPGGVLFDSLGLPHLNPFLHLSQWYNGVGNMVRVIMLAGNAFTA 120

Qy 117 GKIIFAAVPPNFPPTGLSPSQVTMPPHIIVDVQRLEPVLIPDPVRNPNFYHYNQNSDSTI 176
 Db 121 GKIIIVSCIPPFGSGHNLTAQATLPPHVIADVRLTDPIEVPLEDVRLVFNHNDNRNQTM 180
 Qy 177 KLIALMYTLPLRANNAGDDVFTVSCVLRPSDPDPFIPLVPTVESRTKPTFTVPLITVBE 236
 Db 181 RLVCMLLYTLPLRTGGGTGDSFVVAGRVMTCPSPDFNLFPLVPTVEQKTRPFTLPNLPJSS 240
 Qy 237 MSNSRFPILBLEKLYTPSSAFVVOQNGRCCTTGDVLLGTTQLSAVNICTFRGDVTTHIAGS 296
 Db 241 LNSRAPLPISCMGISPDNVQSVQFQNGRCCTLDGLVGTTPVLSHVAKIRCT-----S 294
 Qy 297 HDYTNLASQNSNYDPTTEIPAPLGTDPFVGKIQGMLTQTTRBDGSTRAHKATVSTGVS 356
 Db 295 NGTVINLTDLGTTPHPPEG-PAPIGFPD-LGGCDWHINMT--QFCHSSQTQYDVTTPD 350
 Qy 357 HFTPKLGSVOYTTDNDPQTGNTKFTPVGVIQ--DGNNHQNEPQ--QWVLPNYSGRGT 412
 Db 351 TEVPHLSIQ-----ANGIGSNY-----IGVLSWVSPSPSHSGSQVDLWKIPNYSGIT 400
 Qy 413 HNVHLAPAVAPTFPGEQLLFFRSTMPGCGYPNMMLDCLLPQEWVQHFQCEAAPAQSDVA 472
 Db 401 EATHLAPSVYPPGFGFVLVFFWMSKIPGAY---SLPCLLPQEIYISHLASEQAPTVGEAA 457
 Qy 473 LLRFVNPDTGRVLFECKLHKSGYVTV-----AHTGPHDLVIPPNGYFRFDSWVNPQVYTLAP 528
 Db 458 LLHYVDPDTGRTLGEBKAYPDGFLTCVPGNGASSGPOQL--PINGVVFVSVWSRFPQLKP 515
 Qy 529 MGNAGARRAL 539
 Db 516 VGTASSARGRL 526

RESULT 14
 AAR57091
 ID AAR57091 standard; protein; 530 AA.
 XX AC AAR57091;
 XX 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 05-OCT-1994 (first entry)
 XX Small round virus SRSV/KY/89 capsid protein.
 XX pathogen; acute gastroenteritis; food poisoning; seafood contamination;
 KW diagnostic assay; human calicivirus; small round virus; SRSV; KY89;
 KW Norwalk virus; capsid protein.
 XX Small round structured virus.
 XX WO9405700-A2.
 XX 17-MAR-1994.
 XX 07-SEP-1993; 93WO-US008447.
 XX 07-SEP-1992; 92US-00941365.
 XX (BAYU) BAYLOR COLLEGE MEDICINE.
 XX Matson DO, Bates MK, Jiang X, Graham DY;
 XX WPI; 1994-101125/12.
 DR N-PSDB; AAQ56832.
 XX DNA from Norwalk and related viruses - used for preparing prods. for use
 PT in diagnostic assays, detection and vaccines for Norwalk and related
 PT viruses.
 XX Example 7; Fig 13a; 156pp; English.
 XX

CC The known sequence for Norwalk virus was used to obtain the sequence of
 CC other Norwalk-related viruses such as SRSV/KY/89, an agent from a stool
 CC from an outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide
 CC cDNA sequence includes part of the polymerase region and the capsid
 CC region of the genome; the deduced amino acid sequences are AAR57092 and
 CC AAR57091, respectively. Expression of fragments and derivs. of Norwalk-
 CC related viruses permits development of diagnostic assays to detect
 CC antibodies, antigens, viral genetic material or antivirals. (Updated on
 CC 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS
 CC field.)
 XX
 XX

SQ Sequence 530 AA;

Query Match 38.3%; Score 1106.5; DB 2; Length 530;
 Best Local Similarity 44.2%; Pred. No. 8.9e-91;
 Matches 244; Conservative 89; Mismatches 180; Indels 39; Gaps 16;
 QY 1 MKMASNDANPS--DGSTAN--LYPEVN--NEVMALEPVVGAIAAPVAGQNVDPWIRNPF 56
 DB 1 MMASKDATSSVDGASVOLPEVNASDPLAMDVPVAGSSTAVATAGQNVDPIDPIINPF 60
 QY 57 VQAPGGEFTVSPRNAPCEILWAPLGPDLNPLYLHARMYNGYAGGFEVQVILAGNAFTA 116
 DB 61 VQAPQGEFTISPNTTDFDLVSLGPHLNPFLHLSQMYNGWGNMVRIMLAGNAFTA 120
 QY 117 GKIIFAAVPNPTEGLSPSQVTMFPHIIVDVRLQLEPVLPLPDVRRNFYHYNQSDSTI 176
 DB 121 GKIIIVCIPFGSQQLTIAQATLFPFHVIADVRLDPIEVPLEDVRRNVLFRHNRDNRQQT 180
 QY 177 KLIALMYTLPLRANAGDDVFTVSCRVLTRPSDEDFILVPPTVESRTKPTFVPIITVRE 236
 DB 181 RLVCMLYTLPLTGGTGDSFVAVGRVMTCPSPDFNLFVLPPTVEQKTRFTLNLPLSS 240
 QY 237 MNSRPFPIPLEKLYTGPFSSAFVYQPNQGRCTTDGVLGTLTQLSANNICTFRGDVTHIAG 296
 DB 241 LNSRAPLPISGMGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGT-----S 294
 QY 297 HDYTMNLASQNSWSDYTFEIPAPLGTPTDFVGIQGLMILQTTTREDGSTRAHKATVSTGSV 356
 DB 295 NGTVINILTELDGTPFHFEG-PAPIGFPD-LGGCDWHNMT--QFGHSSQTQYDVDTTPD 350
 QY 357 HFTPKLGSVOYTTDNTDFTQNTKFTPVGVIQ--DGNHQNPEQ--QWLVNYSGRGTG 412
 DB 351 TSVPHLGSIQ-----ANGISGNY-----IGVLSWSPSPHSGSQVDLWKIPNYGSSIT 400
 QY 413 HNVHLAPAV-APTFFGQLLFFRMTMPCGSGYPNNILDCILPQEWQHFQCEAQAQSDV 471
 DB 401 EATHLAPSVSPGF-GEVLVFFMSKIRPGGQ---DSLPCLLPQGYISHLASEQAPTVEG 456
 QY 472 ALLRFVNPDTGRVLFECCLKHSGYVTV----AHTGPHDLVIPNPGYFRFDSWVNOFYTLA 527
 DB 457 PLLHYVDPTDRNLGEFKAYPDGLTICVPNGASSGPQQL--PINGVFVFSVWSRFFYQLK 514
 QY 528 PMNGAGRRRAL 539
 DB 515 PVGTASTARGRL 526

RESULT 15

AAB49702

ID AAB49702 standard; protein; 546 AA.

XX AAB49702;

AC AAB49702;

DT 04-APR-2001 (first entry)

XX Small round structured virus protein SEQ ID 3.

DE Small round structured virus; SRSV; food poisoning.

XX Small round structured virus.

OS Small round structured virus.

XX Small round structured virus.

PN WO200079280-A1.

XX 28-DEC-2000.
 PD
 XX 22-JUN-2000; 2000WO-JP004095.
 PF
 XX 22-JUN-1999; 99JP-00175928.
 PR
 XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.
 XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 PI WPI; 2001-080848/09.
 XX N-PSDB; AAF29143.
 DR
 XX Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX
 PS Claim 1; Page 45-47; 84pp; Japanese.
 XX
 CC This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAP20141 -
 CC AAP20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 CC
 XX Sequence 546 AA;
 SQ
 Query Match 39.3%; Score 1105; DB 4; Length 546;
 Best Local Similarity 44.4%; Pred. No. 1.3e-90;
 Matches 248; Conservative 80; Mismatches 194; Indels 36; Gaps 17;
 QY 1 MKMASNDANPS--DGST--ANLPEVNN--EVMALPEVVGAAIAAPVAGQNVDPWIRNPF 56
 DB 1 MMASKADAPQADGASGAGQLPEVNTADPLMEPVAGPTTAVATAGVNMIDPIVYNF 60
 QY 57 VQAPGGEFTVSPRNAPCEILWAPLGPDLNPLYLHARMYNGYAGGFEVQVILAGNAFTA 116
 DB 61 VQAPQGEFTISPNTTDFDLVSLGPHLNPFLHLSQMYNGWGNMVRIMLAGNAFTA 120
 QY 117 GKIIFAAVPNPTEGLSPSQVTMFPHIIVDVRLQLEPVLPLPDVRRNFYHYNQSDSTI 176
 DB 121 GKIIIVCVPPTGSSLTIAQATLFPFHVIADVRLDPIEVPLEDVRRNVLHXTN-DNQPTM 179
 QY 177 KLIALMYTLPLR--ANNAGDDVFTVSCRVLTRPSDFDFILVPPTVESRTKPTFVPIITV 234
 DB 180 RLVCMLYTLPLRTGGSGNSDSFVAVGRVLTAPSDSFSLFVLPPTIEQKTRFTVPIINPL 239
 QY 235 EMSNSRPFPIPLEKLYTGPFSSAFVYQPNQGRCTTDGVLGTLTQLSANNICTFRGDVTHIA 294
 DB 240 QTLNSRPFSLIQGLMILSPDASQVYQFQNGRCLIDQLGTLTTPATSGQLFVRGKINGA 299
 QY 295 GSHDYTMNLASQNSWSDYTFEIPAPLGTPTDFVGIQGLMILQTTTREDGS--TRAHKATV 351
 DB 300 ----RTLNLTEDVKGFPFAFDS-PAPVGFDF-GKCDWHMRISKTPNNTSSGDPMRVSVS 353
 QY 352 STGSHFTPKLGSVOYTTDNTDFTQNTKFTPVGVIQGLMILQTTTREDGS--TRAHKATV 407
 DB 354 QTNVQGVFPHLGSIQDFEVFNH--PTGDY-----IGTIEWISQSPPTPGTDTNLWEIPDY 406
 QY 408 SGRTHNVHLAPAVAPTFFGQLLFFRMTMPCGSGYPNNILDCILPQEWQHFQCEAQA 465
 DB 407 GSSLQAANLAPVPFPFGGEALVTFVSAFPQNNRSPN-DVPCLLPQGYITHEVSEQA 465
 QY 466 PAQSDVALLRFVNPDTGRVLFECCLKHSGYVTV----AHTGPHDLVIPNPGYFRFDSWV 521
 DB 466 PTWGDAALLHYVDPTDRNLGEFKLPGGYLTCVPNGVAGPQQL--PLNGVFLFVSWVS 523
 QY 522 QFTTLAPMNGAGRRRAL 539
 DB 515 PVGTASTARGRL 526

Db 524 RPYQLKPVGTASTARSL 541

Search completed: June 1, 2004, 13:45:52
Job time : 48.3034 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 12.8419 Seconds
(without alignments)
2166.837 Million cell updates/sec

Title: US-09-926-799-5
Perfect score: 2886
Sequence: 1 MMASNDANPDSGTANLVP.....VNQFTYLPAGMGAGRRRAL 539

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A-COMB.pap.*
2: /cgn2_6/ptodata/2/iaa/5B-COMB.pap.*
3: /cgn2_6/ptodata/2/iaa/6A-COMB.pap.*
4: /cgn2_6/ptodata/2/iaa/6B-COMB.pap.*
5: /cgn2_6/ptodata/2/iaa/PCTUS-COMB.pap.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1161.5	40.2	530	US-08-486-049-3	Sequence 3, Appli
2	322	11.2	669	US-09-617-594A-2	Sequence 2, Appli
3	303	10.5	623	US-09-590-020-4	Sequence 4, Appli
4	300	10.4	623	US-09-590-020-2	Sequence 2, Appli
5	299.5	10.4	626	US-09-590-020-7	Sequence 7, Appli
6	298	10.3	622	US-09-590-020-6	Sequence 6, Appli
7	297.5	10.3	668	US-09-617-594A-4	Sequence 4, Appli
8	127	4.4	2318	US-09-091-219-24	Sequence 24, Appli
9	127	4.4	2318	US-09-660-541-24	Sequence 24, Appli
10	113	3.9	1042	US-09-252-991A-30444	Sequence 30444, A
11	111.5	3.9	776	US-08-961-083-86	Sequence 86, Appli
12	111.5	3.9	776	US-09-536-784-86	Sequence 86, Appli
13	109	3.8	607	US-08-556-422A-4	Sequence 4, Appli
14	105.5	3.7	826	US-09-328-352-7515	Sequence 7515, Ap
15	104.5	3.6	1043	US-08-928-361B-30	Sequence 30, Appli
16	104.5	3.6	1721	US-08-700-651-5	Sequence 5, Appli
17	104.5	3.6	1721	US-08-928-361B-6	Sequence 6, Appli
18	104.5	3.6	1721	US-09-588-995A-6	Sequence 6, Appli
19	102.5	3.6	3892	US-09-328-352-5503	Sequence 5503, Ap
20	101	3.5	667	US-09-094-557-1	Sequence 1, Appli
21	99.5	3.4	558	US-09-071-035-258	Sequence 258, App
22	99.5	3.4	1638	US-09-071-035-258	Sequence 258, App
23	99.5	3.4	1638	US-09-071-035-258	Sequence 262, App
24	99.5	3.4	1638	US-09-071-035-258	Sequence 266, App
25	99.5	3.4	1747	US-09-134-000C-5999	Sequence 5999, Ap
26	98.5	3.4	382	US-09-328-352-4587	Sequence 4587, Ap
27	98.5	3.4	915	US-08-328-322-5	Sequence 5, Appli

28	98.5	3.4	1013	3	US-09-415-522-8	Sequence 8, Appli
29	98.5	3.4	2736	4	US-09-252-991A-30227	Sequence 30227, A
30	98.5	3.4	2972	4	US-09-579-181-2	Sequence 2, Appli
31	98.5	3.4	3118	4	US-09-579-181-1	Sequence 1, Appli
32	98	3.4	627	4	US-09-328-352-7547	Sequence 7547, Ap
33	97.5	3.4	1042	3	US-08-928-361B-11	Sequence 11, Appli
34	97.5	3.4	1042	4	US-09-588-995A-11	Sequence 11, Appli
35	97.5	3.4	1837	3	US-08-928-361B-5	Sequence 5, Appli
36	97.5	3.4	1837	4	US-09-588-995A-5	Sequence 5, Appli
37	96	3.3	1075	5	PCT-US94-07297-41	Sequence 41, Appli
38	95.5	3.3	1091	6	5516630-2	Patent No. 5516630
39	95.5	3.3	1810	4	US-08-793-273C-4	Sequence 4, Appli
40	95.5	3.3	1810	5	PCT-US95-11684-4	Sequence 4, Appli
41	95.5	3.3	2227	3	US-08-475-886-2	Sequence 2, Appli
42	95.5	3.3	2227	3	US-08-475-886-4	Sequence 4, Appli
43	95.5	3.3	2227	3	US-08-475-886-6	Sequence 6, Appli
44	95.5	3.3	2227	3	US-08-397-232-2	Sequence 2, Appli
45	95.5	3.3	2227	3	US-08-397-232-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-486-049-3
; Sequence 3, Application US/08486049
; Patent No. 6572862
; GENERAL INFORMATION:
; APPLICANT: Bates, Mary K
; APPLICANT: Jiang, Xi
; APPLICANT: Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Ave., N.W.
; CITY: Washington, D.C.
; STATE:
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,049
; FILING DATE: June 7, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311.023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-486-049-3

Query Match 40.2%; Score 1161.5; DB 4; Length 530;
Best Local Similarity 45.4%; Pred. NO. 5.6e-104;
Matches 250; Conservative 83; Mismatches 181; Indels 37; Gaps 14;
QY 1 MMASNDANPS-DGST--ANLVPEVN-NEVMALPEVVGAAIAAPVAGQONVIDPWRNF 56
DB 1 MMASKDATSVDGASGAGQLVPEYNASDPLAMDPEVAGSTAVATAGQVNPDPWIIINF 60

QY 57 VOAPGEBTVSPNAPGELLSAPLGPDLNPNYLHARMYNGYAGGEVQVVLGNAGFTA 116
DB 61 VOAPGEBTVSPNAPGELLSAPLGPDLNPNYLHARMYNGYAGGEVQVVLGNAGFTA 120
QY 117 GKIIIFAAVPPNPPTGLSPQVTPMPPHIIIDVRQLPVLIIPLDPVNNFYHYNQSNDSIT 176
DB 121 GKIIIVSCITPPGGSHNLITIAQATLFPHVIADVRLDPIEVLDPVNNFYHYNQSNDSIT 180
QY 177 KLIAMLYPLRANAGDDVFTVSCVRLTRPSPDPDFILVPTVESRTKPTVPLTVTEE 236
DB 181 RLVCMLYPLRANAGDDVFTVSCVRLTRPSPDPDFILVPTVESRTKPTVPLTVTEE 240
QY 237 MNSRPPIDLEKLYTGPSSAFVQVQNGRCTTGDVLLGTTQLSAVNICTRFGDVTHIAGS 296
DB 241 LNSRAPLPISSMGISPDVNVQVQNGRCTTGDVLLGTTQLSAVNICTRFGDVTHIAGS 294
QY 297 HDVTNMLASQNSNYDPTTEIPAPLGTDFVKGIOGMLTQTTREDSGSTRAHKATYSTGSV 356
DB 295 NGTVINLTDLGTPPHFPEG-PAPIGFPD-LGGCDWHINMT--QFCHSSQTQYDVTTPD 350
QY 357 HPTPKLGSVQYTTDNDTQNTKFTVGVQI--DGNHONHOPE--QWVLPNYSGRGT 412
DB 351 TFVPHLSIQ-----ANGISGNY-----VGLSWISPSHSGSVDLWKIPNYGSSIT 400
QY 413 HNVHLAPAVAPTPFGGQLLFFRSTMPGCGYFNMNLDCLLPQEWVQHFCEAAPQSDVA 472
DB 401 EATHLAPSVYPPGFGVGVFFNFKMFGPGAY---NLPCLLPQEIYSHLASEQAPTVEAA 457
QY 473 LLRFVNPDTGRVLFECMLKHSQVTV-----AHTGPHDLVIPPNGYERFDSWNOFTLAP 528
DB 458 LLHVDPDTGRNLGEKAYPDGLTFCVNGASSGPOOL--PINGVFFVSWSRFQLKP 515
QY 529 MNGAGRRAL 539
DB 516 VGTASSARGRL 526

RESULT 2
US-09-617-594A-2
; Sequence 2, Application US/09617594A
; Patent No. 6541458
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.1
; CURRENT APPLICATION NUMBER: US/09/617,594A
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-617-594A-2

Query Match 11.2%; Score 322; DB 4; Length 669;
Best Local Similarity 25.4%; Pred. No. 3.1e-22;
Matches 125; Conservative 69; Mismatches 158; Indels 140; Gaps 22;
QY 11 SDGSTANLVPEVNNEV--MALEPVVGAATAAPVAGQNVDPWIRNPNVQAPGGEFTV-- 66
DB 125 AUGDSITTEQGTGVLGGVIAEPESAQMAATAADATKSVDSW--ESF-----FSPHT 175
QY 67 ----SPRNAPGELLSAPLGPDLNPNYLHARMYNGYAGGEVQVVLGNAGFTAAGKIIFA 122
DB 176 SVNWSTSETQKILFKQSLGPLNPLYLHSLKLYVAVSGSDVDRFSGSGVFGKLAIAI 235

QY 123 AVPPNPPTGLSPSQVT---MEPHLIIDVRQLPVLIIPLDPVNNFYHYNQSNDSITKLI 179
DB 236 VVPP-----GVDFVQSTSMQLQPHVLFDAQVEFVIFSPDLRSTLYHLMSDITDT-SLV 289
QY 180 AMLYTPLRANAGDDVFTVSC--RVLTSPSPDPDFILVPP-----TVES----- 222
DB 290 IMVNDL-INPYANDNSNGCIVTVETKPGDPFKHLLKPPGSMMLTHGSIPLDKSS 348
QY 223 -----RTKPTVPLITVEEMNSRPP 243
DB 349 LWIGNRYWSDITDFVIRPFVQANRHFDFNQETAGWSTPRFRPITI--TISENGS---- 402
QY 244 IPLEKLYTGPSSAFVQV--PQNGRCTTGDVLLGTTQLSAVNICTRFGDVTHIAGSDHYTM 301
DB 403 -----KLGTVATDIYIVPGIPDPGWDPTTIGEE--TPAGDYSITNGSGN---- 444
QY 302 NLASQNSNYDPTTEIPAPLGTDFVKG-IOGMLTQ-----TTREDGSTRA 346
DB 445 DIATAN--AYDSADVI---TNTTFRGMVTCGALORAWGDKKISSTAFITTAKEGNTLK 499
QY 347 HKATVSTGSHVHTPKLGSVQYTTDNDTQNTKFTVGVQI--DGNHONHOPE--QWVLPNYSGRGT 412
DB 500 PSNTIDMTKI-----AVYQDTHVGRDVQTSDDTLAILGYTGIGEAIGSNRDSVVRIS 552
QY 403 VLPNYSGRGTGHN 414
DB 553 MLPETGARGGNH 564

RESULT 3
US-09-590-020-4
; Sequence 4, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-4

Query Match 10.5%; Score 303; DB 4; Length 623;
Best Local Similarity 26.1%; Pred. No. 1.9e-20;
Matches 128; Conservative 55; Mismatches 167; Indels 140; Gaps 23;
QY 12 DGSTANLVPEVNNEV--MALEPVVGAATAAPVAGQNVDPWIRNPNVQAPGGEFTV-- 66
DB 82 DGSIT--TPEQGTMTGGVIAEPAQMSAAADATKSVDSW-----EAFSPHTSVNM 133
QY 67 SPNAPGELLSAPLGPDLNPNYLHARMYNGYAGGEVQVVLGNAGFTAAGKIIFA 126
DB 134 STSETQKILFKQSLGPLNPLYLHSLKLYVAVSGSDVDRFSGSGVFGKLAIAIVVP 193
QY 127 NFPTGLSPSQVT---MEPHIIIDVRQLPVLIIPLDPVNNFYHYNQSNDSITKLI 183
DB 194 -----GIEPVQSTMLQYPHVLFDAQVEFVIFSPDLRSTLYHLMSDITDT-SLVIMVY 247
QY 184 TPLRANAGDDVFTVSC--RVLTSPSPDPDFILVPP-----TVES----- 222
DB 248 NDL-INPYANDTNSGCIIVTVETKPGDPFKHLLKPPGSMMLTHGSIPLDKSS 306


```
QY 223 -----RTKPTVPILTVEEMSNRRPIPLE 247
Db 307 NRHWSIDTPIIRPFVQANRHFDFNQETAGMSTRFRPRTI---TVSE-SN-----MS 356
QY 248 KLYTGPSSAFVQ--PONGRCCTDGLTQLSAVNICTFRG-DVTHIAGSHDYTMNLA 304
Db 357 KLGIGVATDIYVPGIPDGWPTTIEQL--TPAGIYSITASNGTDITTAAG-----405
QY 305 SONWNSYDTEIPAPLPTPDF-----VGKIQ-----GMLTQTTREDGGSTRAHK 348
Db 406 -----YDAAEI---VNTTFKSMYICGLQRAWGDKKISNTAFITTAVRKGNSTEPSN 456
QY 349 ATVSGSVHFTPKLGSVQVTTDNDPQONT---KFTPGVLODGNHQNHPQWVL 404
Db 457 TIDMTKL-----VYODAHVGEVQTSITLALLGYTGIGEBAGSDRDKVVRISVL 508
QY 405 PNYSGRTGHN 414
Db 509 PETGARGGNH 518

RESULT 4
US-09-590-020-2
; Sequence 2, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; PRIOR FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-2

Query Match 10.4%; Score 300; DB 4; Length 623;
Best Local Similarity 26.9%; Pred. No. 3.7e-20;
Matches 97; Conservative 45; Mismatches 113; Indels 106; Gaps 14;

QY 12 DGSITANLVPVENEV--MALEPVVGAATAAPVAGQONVIDPWIRNFFVQAPGGEFTV--- 66
Db 82 DGSIT--TPEQGTWVGVIAPESQAQMSAADMATKGSVDSEW-----EAFPSFHTSVNM 133
QY 67 SPRNAPGELLWSAPLGPDLNPLYSLARMYAGGFEVQVILAGNAPTAGKIIFAAVPP 126
Db 134 STSETQKILFKQSLGPLNPLYSLAKLYVAVSGSIEVRFSIGSGVFGGKLAIVVPP 193
QY 127 NPFTGLSPSQVT---MFPPIIIVDRQLEPVLIPLDVNRNFFHYNQSNDSITKLIAMLY 183
Db 194 -----GIEPVQSTMLQYPHVLFDAQVEPVIFALPDLRSNLYHLMSDITTT-SLVIWY 247
QY 184 TPLRANNAGDDVFTVSC--RVLTRSPDPDFILVPPPTVESRTKPTVPILTVEEMSNR 241
Db 248 NDL-INPYANDTNSSGCIIVTETKPGDPFKHLLKPP--GSMITHGVSVDLIPKSS-- 302
QY 242 PPIPLEKLYTGPSSAFVQVQ---QNGR-----CITDGLVLTGTTQLSAVNICTFRGDVTHI 293
Db 303 ---LWIGNRHSIDITDIIRPFVQANRHFDFNQETAGMSTRFRPRTIITSONGAKUGI 360
QY 294 AGSHDYTMNLSAQNSNYD-PTEEIPAPLGTDPFVGKIQGMLTQTTREDGSTR--HKAT 350
Db 361 GVATDIYVPGIPDGWPTTIEQLIPA-----GDYA-----ITNGTGNDITATGYDTRAD 410
QY 351 VSTGSHVF-----TPKLGVSQVYTT-----DTNNDPQTQNTKFTPVGVITQ 390
Db 411 IIKNTNFRGMWYICGLQRAWGDKKISNTAFITATLGDNNNKINPCNTIDQSKIIVVQ 470
QY 391 D---GNHQNHPQWVLNYSQ-----RTCHNVHLAPAVAPTTP-----GEQLLPRSTM 437
Db 471 DAHVKKQAQTSDDTLALLGYTGIGESQAIQSORDRVRIS-TLPETGARGGNHPIFYKSI 529
QY 438 PGCSYPPNNMLDCLLPQEWQHFCEAA-----PAQSDVALLRFPVNPDTGRVLPFCKLH 491
Db 530 K--LGVTVRSIDVFNQ--ILHTRQLSUNHYLLPDPDS-FAYVRII--DSNGSWFDIGID 582
QY 492 KSGYVTVVAHTGPHDLVIP 509
Db 583 SDGFSFVGSGFGKLEFP 600

RESULT 6
```

```
QY 223 -----RTKPTVPILTVEEMSNRRPIPLE 247
Db 307 NRHWSIDTPIIRPFVQANRHFDFNQETAGMSTRFRPRTI---TVSE-SN-----MS 356
QY 248 KLYTGPSSAFVQ--PONGRCCTDGLTQLSAVNICTFRG-DVTHIAGSHDYTMNLA 304
Db 357 KLGIGVATDIYVPGIPDGWPTTIEQL--TPAGIYSITASNGTDITTAAG-----405
QY 305 SONWNSYDTEIPAPLPTPDF-----VGKIQ-----GMLTQTTREDGGSTRAHK 348
Db 406 -----YDAAEI---VNTTFKSMYICGLQRAWGDKKISNTAFITTAVRKGNSTEPSN 456
QY 349 ATVSGSVHFTPKLGSVQVTTDNDPQONT---KFTPGVLODGNHQNHPQWVL 404
Db 457 TIDMTKL-----VYODAHVGEVQTSITLALLGYTGIGEBAGSDRDKVVRISVL 508
QY 405 PNYSGRTGHN 414
Db 509 PETGARGGNH 518

RESULT 4
US-09-590-020-2
; Sequence 2, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; PRIOR FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-2

Query Match 10.4%; Score 300; DB 4; Length 623;
Best Local Similarity 26.9%; Pred. No. 3.7e-20;
Matches 97; Conservative 45; Mismatches 113; Indels 106; Gaps 14;

QY 12 DGSITANLVPVENEV--MALEPVVGAATAAPVAGQONVIDPWIRNFFVQAPGGEFTV--- 66
Db 82 DGSIT--TPEQGTWVGVIAPESQAQMSAADMATKGSVDSEW-----EAFPSFHTSVNM 133
QY 67 SPRNAPGELLWSAPLGPDLNPLYSLARMYAGGFEVQVILAGNAPTAGKIIFAAVPP 126
Db 134 STSETQKILFKQSLGPLNPLYSLAKLYVAVSGSIEVRFSIGSGVFGGKLAIVVPP 193
QY 127 NPFTGLSPSQVT---MFPPIIIVDRQLEPVLIPLDVNRNFFHYNQSNDSITKLIAMLY 183
Db 194 -----GIEPVQSTMLQYPHVLFDAQVEPVIFALPDLRSNLYHLMSDITTT-SLVIWY 247
QY 184 TPLRANNAGDDVFTVSC--RVLTRSPDPDFILVPPPTVESRTKPTVPILTVEEMSNR 241
Db 248 NDL-INPYANDTNSSGCIIVTETKPGDPFKHLLKPP--GSMITHGVSVDLIPKSS-- 302
QY 242 PPIPLEKLYTGPSSAFVQVQ---QNGR-----CITDGLVLTGTTQLSAVNICTFRGDVTHI 293
Db 303 ---LWIGNRHSIDITDIIRPFVQANRHFDFNQETAGMSTRFRPRTIITSONGAKUGI 360
QY 294 AGSHDYTMNLSAQNSNYD-PTEEIPAPLGTDPFVGKIQGMLTQTTREDGSTR--HKAT 350
Db 361 GVATDIYVPGIPDGWPTTIEQLIPA-----GDYA-----ITNGTGNDITATGYDTRAD 410
QY 351 VSTGSHVF-----TPKLGVSQVYTT-----DTNNDPQTQNTKFTPVGVITQ 390
Db 411 IIKNTNFRGMWYICGLQRAWGDKKISNTAFITATLGDNNNKINPCNTIDQSKIIVVQ 470
QY 391 D---GNHQNHPQWVLNYSQ-----RTCHNVHLAPAVAPTTP-----GEQLLPRSTM 437
Db 471 DAHVKKQAQTSDDTLALLGYTGIGESQAIQSORDRVRIS-TLPETGARGGNHPIFYKSI 529
QY 438 PGCSYPPNNMLDCLLPQEWQHFCEAA-----PAQSDVALLRFPVNPDTGRVLPFCKLH 491
Db 530 K--LGVTVRSIDVFNQ--ILHTRQLSUNHYLLPDPDS-FAYVRII--DSNGSWFDIGID 582
QY 492 KSGYVTVVAHTGPHDLVIP 509
Db 583 SDGFSFVGSGFGKLEFP 600

RESULT 6
```

US-09-590-020-6
; Sequence 6, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; PRIOR FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Feline calicivirus
; US-09-590-020-6

Query Match 10.3%; Score 298; DB 4; Length 622;
Best Local Similarity 27.1%; Pred. No. 5.8e-20;
Matches 98; Conservative 43; Mismatches 114; Indels 106; Gaps 14;

Qy 12 DGTANLVPEVNNVEV--MALEPVVGAIAAPVAGQONVIDPWRNNFVQAPGGFTV--- 66
Db 82 DGSIT--TPBQGTWVGVIAPSAQMSAADMATGKSVSEW-----EAFSFTSVNW 133

Qy 67 SPRNAPGEIILWSAPLGDPLNPLYSLHARMYNGVAGFEVQVILAGNAFTAGKIIFAAVPP 126
Db 134 STSETQKILFKQSLGSLPLNPLYSLHAKLVVALAGSVFVRSISGSGVFGKLAIVVPP 193

Qy 127 NFPTGSLSPQV---MFPHIIIVDVRQLEPVLPLPDVRNNFYHYNQSDSTIKLIAMLY 183
Db 194 -----GIEPVQSTSMLOQPHVLFDAQVPEVFIPTDLRNSLYHLMSDITDTT--SLVIMY 247

Qy 184 TPLRANNAGDDVFTVSC--RVLTSPSPDFDFIPLVPP-----TVES----- 222
Db 248 ND-L-INYANTNSGGCIVTETKPGDFKHLKPPGSMLTRGSPVSDILIPKSSSLWIG 306

Qy 223 -----RHKPFTVPLTVEEMSNRFRPIPLE 247
Db 307 NRHWSIDTDFIIRPFVQANRHFDPNQBETAGMSTFRPRITI-----TVSESNMKGILGVA 363

Qy 248 KLY-----TGPSAFVVPQNGRCITTDG-----VLLGTTQLSANNI 283
Db 364 TDYIVPGIDGWPDTTIEQLTPAGIYSITASNGTDITTAAGYDAAETIVNTTNFKSMYI 423

Qy 284 C 284
Db 424 C 424

RESULT 7
US-09-617-594A-4
; Sequence 4, Application US/09617594A
; Patent No. 6541458
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V
; FILE REFERENCE: 454313-3151.1
; CURRENT APPLICATION NUMBER: US/09/617,594A
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
; US-09-617-594A-4

Query Match 10.3%; Score 297.5; DB 4; Length 668;
Best Local Similarity 26.8%; Pred. No. 7.3e-20;
Matches 110; Conservative 52; Mismatches 147; Indels 101; Gaps 20;

Qy 12 DGTANLVPEVNNVEVMALEPVVGAIAAPVAGQONVIDPWRNNFVQAPGGFE----- 64
Db 127 DGSVT--TPEQGT-----LVGGVIAEPNA-QMSAVADVATGKSVDSWEAFPSFHTSV 176

Qy 65 TVSPRNAPGEIILWSAPLGDPLNPLYSLHARMYNGVAGFEVQVILAGNAFTAGKIIFAAV 124
Db 177 NWSTSETQKILFKQSLGSLPLNPLYSLHAKLVVANGSGIEVFRSISGSGVFGKLAIVV 236

Qy 125 PPNFTEGLSPSOVT---MFPHIIIVDVRQLEPVLPLPDVRNNFYHYNQSDSTIKLIAM 181
Db 237 PP-----GIDPVOSTSMLOQPHVLFDAQVPEVFIPTDLRNSLYHLMSDITDTT--SLVIM 290

Qy 182 LYTPLRANNAGDDVFTVSC--RVLTSPSPDFDFIPLVPPVETSRTPKFTVPILTVEEMSN 239
Db 291 IYNDL-INYANDNSGGCIVTETKPGDFDFHLLKPPG-----SMLTHGSIFS 339

Qy 240 SRPIPLEKLYTGPSSAFVVPQNGRCITTDGVLGTTQLSANNICTFRGDVTHIAGSHDY 299
Db 340 DLIP-----KSSSLWIGNRHWSIDITDFV-----IKPFVFO-----ANRH-F 374

Qy 300 TMNLASQWNSNYDPTEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGVS--- 356
Db 375 DFNQETAGWS-----TPRF-----RPITITVSEKGGSKL-GIGVATDSIVPG 415

Qy 357 -----HFTPKLGSVQVTTDNDFTQNTKFTFVGVQIDGNN 394
Db 416 IPDGMPTTPEIKLTP-AGDYAITNGGNDITTAAD--YDGASIIKNTN 462

RESULT 8
US-09-091-219-24
; Sequence 24, Application US/09091219
; Patent No. 6171592
; GENERAL INFORMATION:
; APPLICANT: STUDDERT, Michael J.
; APPLICANT: CRABE, Brendan S.
; APPLICANT: PENG, Li
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
; FILE REFERENCE: 040268/0151
; CURRENT APPLICATION NUMBER: US/09/091,219
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: PCT/AU96/00815
; EARLIER FILING DATE: 1996-12-18
; EARLIER APPLICATION NUMBER: AU PN7201
; EARLIER FILING DATE: 1995-12-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 2318
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
; US-09-091-219-24

Query Match 4.4%; Score 127; DB 3; Length 2318;
Best Local Similarity 22.3%; Pred. No. 0.021;
Matches 73; Conservative 39; Mismatches 134; Indels 82; Gaps 14;

Qy 92 LARMYNGYAGGFEVQVILAGNAFTAGKIIFAAVPPNFPTEGLSPSOVTMFPHIIIVDVRQL 151
Db 380 LTDSYAYMRNGWDVEVTAVGNQNGCLLVAMVPYLSIQKRELYQLTLFPHQFINPRTN 439
Qy 152 EPVLIPLDPVRNNFY-HYNQSDSTIKLIAMLYTPLRANNAGDDVFTVSCRVLTPSPDF 210

Db 440 MTAHTVPPGVNRYDQYKVKHPWT--LVVMVAPLTVNTEG-----APQI 483
QY 211 DFILVPPVESRTPFTVPILTVEMSNRRPIPLEKLY-----TGPSSAFV-----259
Db 484 KYVANIAPT-----NVHAGEFPPSKEGIFPVACSDGYGLVTTDPKTDADPVYGVKVFN 535
QY 260 QPON--GRCITDGVLLGTQTSANVNICTFRGDVTTHAGSHD-----YTNLASQNW 309
Db 536 PPRNQLPGRFTN---LLDVAE-ACPTFLRFEFGVPYVTKTDSRDLVAQFDMSLAOKMS 591
QY 310 NYDPTTEEIPAPLGTDFVCKIQGMLTQTTREDGSTRAHKATVSTG-----SVHFTPK 361
Db 592 N-----TFLAGLAQYTYQYSGTINLH--FMFTGPTDAKARYMWAYAPP 632
QY 362 LGSVQYTTDT-----NNDFQTGONTKFT 384
Db 633 GMEPPKTPAAAHCIAHAEWDTGLNSKFT 660

RESULT 9

US-09-660-541-24
; Sequence 24, Application US/09660541
; Patent No. 6531136
; GENERAL INFORMATION:
; APPLICANT: STUDDERT, Michael J.
; APPLICANT: CRABB, Brendan S.
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
; FILE REFERENCE: 040268/0151
; CURRENT APPLICATION NUMBER: US/09/660,541
; CURRENT FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/091,219
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: AU PN7201
; PRIOR FILING DATE: EARLIER FILING DATE: 1995-12-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 2318
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
US-09-660-541-24

Query Match 4.4%; Score 127; DB 4; Length 2318;
Best Local Similarity 22.3%; Pred. No. 0.021;
Matches 73; Conservative 39; Mismatches 134; Indels 82; Gaps 14;
QY 92 LARMYNGAGGFEVQVILAGNAFTAGKIIFAAVPPNFPTEGLSPSQVTMFPHIIVDVRL 151
Db 380 LTDSYAYMRNGWDVEVTAAGNFGCGLLVAMVPYLSIQKRELYQLTLFPHQINPRTN 439
QY 152 EPVLIPLDPVRNFFY-HYNQSNDSITIKLIFAMLYPLRANNAGDDVFTVSCRVLTRPSPDF 210
Db 440 MTAHTVPPGVNRYDQYKVKHPWT--LVVMVAPLTVNTEG-----APQI 483
QY 211 DFILVPPVESRTPFTVPILTVEMSNRRPIPLEKLY-----TGPSSAFV-----259
Db 484 KYVANIAPT-----NVHAGEFPPSKEGIFPVACSDGYGLVTTDPKTDADPVYGVKVFN 535
QY 260 QPON--GRCITDGVLLGTQTSANVNICTFRGDVTTHAGSHD-----YTNLASQNW 309
Db 536 PPRNQLPGRFTN---LLDVAE-ACPTFLRFEFGVPYVTKTDSRDLVAQFDMSLAOKMS 591
QY 310 NYDPTTEEIPAPLGTDFVCKIQGMLTQTTREDGSTRAHKATVSTG-----SVHFTPK 361
Db 592 N-----TFLAGLAQYTYQYSGTINLH--FMFTGPTDAKARYMWAYAPP 632
QY 362 LGSVQYTTDT-----NNDFQTGONTKFT 384
Db 633 GMEPPKTPAAAHCIAHAEWDTGLNSKFT 660

RESULT 10

US-09-926-799-5.ra1

US-09-252-991A-30444
; Sequence 30444, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30444
; LENGTH: 1042
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30444

Query Match 3.9%; Score 113; DB 4; Length 1042;
Best Local Similarity 22.8%; Pred. No. 0.13;
Matches 126; Conservative 53; Mismatches 234; Indels 140; Gaps 28;
QY 11 SDGSTANLVPEVNEVMALEPVVGAIAAPVAGQONVI-----DPWIRNFVOAPGG 62
Db 21 ADGATESTV-----VLAGPLALPAASSATALITVPLASGVPGVK---VOLPLP 65
QY 63 EFTVSPRNPAGEILLWSAPLGDNLNPLS-HLARMYNGVAGGFEVQVILAGN-AFTAGKII 120
Db 66 SAVVSPIGLPLSPVRITLAPTSAPVITLPLLALITGVAGATVSTVVLGSLVLPAGS-- 123
QY 121 FAAVPPNFPTEGLSPSQVTMFPHIIVDVRLQLEPVLIPLDVRNFFHYHYNQSNDSITIKLJA 180
Db 124 WATALITVPLASGVPGVKQLP---LPSARVWPVIGLPLSVR-----ITRTGSAVPL-- 173
QY 181 MLYTPLRANNAGDDVFTVSCRVLTRPSPDFIFLVPPTVESRTPKFTVPILTVEMSN 240
Db 174 -SLTPLLGSITGVAGATASTVLTGP-----LVLPAASWATALITVPLVSGVPGVKL 224
QY 241 RPIPLEKLYTCPSAFVVPQNGRCTTGDVLGTLTQTSANVNIC--FRGDTVTHIAGSHD 298
Db 225 QLPLP-----SAVACRIGLPPPSVRIT-----VAPASAVPLITVPLLGSSTGVAGATE 272
QY 239 YTMILA-----SQWNSYDPTTEIPAPLGTDFVCKIQGMLTQTTT-----EDGSTRAH 347
Db 273 STVVLGSLVLPPLPAAASALITE--PLASGVPGV--KLQLPLPSAVTWPIGLPPPSVRIT 328
QY 348 KATVSTGSHVFTPKLGSVQYTTDNNDFQTGONTKFTPVGVITQDGNHONHNEPQMWLPNY 407
Db 329 VAPASAVPAITLPLLG-----LTTG-----V 349
QY 408 SGRTGHNVL-APAVAPTFGEQLLFFRSTWPGSCGYNNMLDCLLPQ--EWWQHFQCEA 464
Db 350 AGATASTVVLVGLVPLPAGSWATALI---TVPLASGVPGVKLQLPLPSAVAMPVIGLPPPS 406
QY 465 -----APAQSDVALLRFVNPDGTGRVLFPECKLHKSQYVTVVAHTGPHDLVTPPGYFRFDSW 519
Db 407 VRITVAPASAVPAI---TLPLLGSITGAAGATES---TVVLGAP--LVLPA-----ASW 452
QY 520 VNQFYTLAPMGNG 532
Db 453 ATAL--TTVPLDSDG 464

RESULT 11
US-08-961-083-86
; Sequence 86, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

Db 256 DDK-----VPAVTA-SVSNGLATVVPVREGEPRVRVIAKANGDILGEYRLHFTKOKSL 309
Qy 273 LGTTQLSAVNICTFR--GDV-----THIAGSHDYTNLILASQNSVNDPTTEIPAPLG 322
Db 310 LSHKPAVAAKQARLLQVGALELPTKVPVYFGKDGKETKDLTVEM-----EEVPAENL 363
Qy 323 PTDVFGKIQGMLTQTTRDGGSTRAHKAIVSTGSHFTPKLGSV-----QYTTDTNNDPQT 377
Db 364 TK-----AGQFTVRGRVLGSLVAEIT-----VRVTDKLGELTSDNPNYDENSNOAFAS 412
Qy 378 CON-----TKFTPVGVIQGNHNEPQWVLPVNSGRGTGNVHLAPAVAPFPGLQLFF 433
Db 413 ATNDIDKNSHRVDVINDGDHSEN--RRWT--NWSPTSSNEVSAGV-----IF 458
Qy 434 RSTMPGCSGYPNMNLDCLLPQWVQHFQCEAAPQSDVALLRFVNPDTGRVLFECKLHKS 493
Db 459 REN-----GKIVERTVQGVQFFADSGTDAFSLVRLRYVGPE-----FEVPTYYS 505
Qy 494 GYVTVAHGTGPHDLVTPPN 511
Db 506 NY--QAYDADHPFNPN 521

RESULT 13

US-08-556-422A-4
; Sequence 4, Application US/08556422A
; Patent No. 6576754

GENERAL INFORMATION:

; APPLICANT: HALL, Kathryn T.

; APPLICANT: FREEMAN, Gordon J.

; APPLICANT: SCHULTZE, Joachim L.

; APPLICANT: BOUSSIOTIS, Vassiliki

; APPLICANT: NADLER, Lee M.

; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES

; FILE REFERENCE: DFN-005CPA2

; CURRENT APPLICATION NUMBER: US/08/556,422A

; CURRENT FILING DATE: 1995-11-09

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 607

; TYPE: PRT

; ORGANISM: Mus musculus

US-08-556-422A-4

Query Match 3.8%; Score 109; DB 4; Length 607;

Best Local Similarity 18.9%; Pred. No. 0.13;

Matches 106; Conservative 59; Mismatches 165; Indels 230; Gaps 29;

Qy 52 IRNFEVQAPGGEFTVSPRNAPGEILWSAPLG-----ELLSADADKQCKSPKDKPKDCQNYIKILLPINS 98
Db 39 LNSNLSFLPGGEYQ-----LLTCTAASFPLCAYIHIASTFLAQDEAGNVILEDGKGHCPCFPDNFKSTALV 143
Qy 99 YAGGFEVQVILAGNA-----FT-----AGKIIF-----AAVPPNF-----128
Db 91 H-----LLTCTAASFPLCAYIHIASTFLAQDEAGNVILEDGKGHCPCFPDNFKSTALV 143
Qy 129 -----PTEGLSPQVTFMPPHIIIVDVQLEPVLIPDV 161
Db 144 VDGELYTGVTSFQGNDFPAISRSQSRPTKTESSLNWLQDPAFVASATSPESLGSPIGDD 203
Qy 162 RNNFVHYNQSN-----DSTIKLIAMLYTPLRANNAGDDV-----FTVSCRVLTRSP 208
Db 204 DKYIFFSETGGEFFENTII--VSRVARVCKGDEGGERVLQQRWTSFLKQALLCSRPDD 261
Qy 209 PDFDFILVPPTVESRTKPTTPIILVTEMSNSRFPPIPLEKLYT--GPSSAFVVPQNGRGT 267
Db 262 GF-----PENV-----LQDVFTLNP-----PDWRKT 284
Qy 268 TD-GVLL-----GTTQLSAVNICTF-RGDVTHIAGSHDYTNLILASQNSVNDPTTEIPAP 320
Db 285 LSIQVFTSQWHRGTTEGSA--ICVETMNDVQKAFDGLYKKVNRETQW--YTETHQVPTP 340

Qy 321 -----LGTDPDFVGKIQGMLTQTTRDGGSTRAH-----KATVSTGSV 356
Db 341 RPGACITNSARERKINSSQLQPD---RVNLFLKOHFLMDGOVRSRLLLLOPPARYQORVAV 397
Qy 357 HFTPKLGSVQYTTTNDPQTGONTKFTPVGVIOQGNH-----QN 397
Db 398 HRVFGHLST-----YDVLFLGTG---DGRHLKAVTLSSRVHIIIBELOIFPQG 441
Qy 398 EP-QQWVLPNYSGRGTGNVHLAPAVAPFPGEQLLFFRSTMPGCSGYPNMNLDCLLPOB- 455
Db 442 QPVQNLLLDSHGGLLYASSHSGVQVP-----VANCSLYPTCG-DCILLARDP 487
Qy 456 ---WVQHFCQBAAPQSDVA 472
Db 488 YCAWGTGSACRLASLYQPDLA 507

RESULT 14

US-09-328-352-7515

; Sequence 7515, Application US/09328352

; Patent No. 6562958

GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 7515

; LENGTH: 826

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-7515

Query Match 3.7%; Score 105.5; DB 4; Length 826;

Best Local Similarity 22.7%; Pred. No. 0.47;

Matches 116; Conservative 60; Mismatches 180; Indels 155; Gaps 32;

Qy 58 QAPGGEFTVSPRNAPGEILWSAPLGPDLPNLYSLHARMYGVAGGFEVQVILAGNAFTAG 117
Db 359 KAPRFEVTSAPTIACTTIVGSRIDNV-----AADMPPGVIRGYDV-----ITG 403
Qy 118 KIIFAAVP-----PNP-----PTEGLSPSQVTFMPPHIIIVDVQLEPVLIPL-----PDV-----R 162
Db 404 KLRWAFDPNPDPNVYLKPGEIYKRSSANSWAAMSYD--PQMTVFLPMGSSSVVDIWGNR 462
Qy 163 NNFYH-YNQS---NDSTIKLIAMLYTPLRANNAGDDVFTVSCRVLTRPSDPDFIF---L 215
Db 463 NPLDHKYNSTVLALDATTGKEKWVYQTVH-----NDLW-----DFDLPMPQSL 505
Qy 216 VP-PTVESRTKPTTV-----PILTVEMSNSRFPPILEKL-YTGPSA 256
Db 506 VDFFLKOGTTKPAVVGITKSCQFFVLDRVTKPLTKVIEQPVKAANIPEGEYSLTQPRS- 564
Qy 257 FVWQPQNGRCT--TDGVLLGTQ---LSAVNICTPRGDVTIAGSHDYTM-----NLASON 307
Db 565 -VEMPOIGNOTLTESDMKGATFPDQLMCRINFKSNRYEGLTAPCTDVSLSPPGSLGGNN 623
Qy 308 WSN---YDPTBEIPAPLGPDPVFGKIQ-QM---LQTTRDGGSTRA---HKATVSTGSHVP 358
Db 624 WGSTAFDPTHRVM-----FVNDMLGLMWQLIKQTPEDIKIQASGCKVKNTGMGAV-- 674
Qy 359 TPKLGS-----VQYTTTNDPQTGONTKFTPVGVIOQGNHNEP 399
Db 675 -PMKGTPIYVKNKRFMSALSIPCKPPFTMTAIDMKTRQVAVQVPLGTVEB-----725
Qy 400 QQWVLEPNYSGRGTGNVHL-APAVAPTFPGEQ-----LLFFRSTMPGCSGYPNMNLDCLLP 453
Db 726 -----TGPLGLKMLKAPIGHPTTGGPMATOGGLVFFAAT-----QDYILRAEDSSTG 773
Qy 454 QE-WVQHFCQBAAPQSDVALLRFVNPDTGR 483

us-09-926-799-5.ra1

Wed Jun 2 09:13:29 2004

Qy 242 FEIPLEKLYTGPSAFVVPQNGRCTTGDVLLGTTQLSAVNICTFRGDVTHIAGSHDYTM 301
Db 303 ---PTTGLPFNPPPTGHLINPTNNN-TMD-----SSFAGAYKAV 337
Qy 302 NLASQWNSNYDPTTEIPAPLGTDFPVGKIQGM-----LTOTRE--DGSTRAHKAT 350
Db 338 -----SNGIKTDNV---YGLP--VGEITGLPKDPGSDIPFNSTTGELVDFSTGKPINN 385
Qy 351 VSTGSVHFTPKLGSVOYTTDNDNDFQTGONTKFTPVGVIQDGNHONBPQQ----- 401
Db 386 STAGIVSGRKLPLPIE--DENGNLFPDSTN---LPI-----DGNOLVNPETNSTVSGSTS 436
Qy 402 -----WVLPN-----YSGRTGHNH 416
Db 437 GTTKPKGIPVNGGVVPEEAKQADKKGGLIVPPTNSINKDPVTNTQYSNTTGNII 496
Qy 417 LAPAVAPTFFGEQLLFSTRMTGCGSYPMNLDCLLPQEWVQHFQCEAAPAQSDVALLRF 476
Db 497 --PETGKVIPTG-----SLPGLNYPSPFN---TPQQ-----TDEITGKPVDTVTGLP 537
Qy 477 VNPDTGEVL 485
Db 538 YDPSTGEII 546

Search completed: June 1, 2004, 13:57:58
Job time : 13.8419 secs

Db 774 KELW-----KSRMPVGSOGTPIISVSPKTKG 799
RESULT 15
US-08-928-361B-30
; Sequence 30, Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: SPECIES INFECTIONS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSES: PETERS, VERNY, JONES & BIKSA
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-1840
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,062
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Verny, Hana
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480.76-1 (HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-1677
; TELEFAX: 650-324-1678
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1043 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-928-361B-30

Query Match 3.6%; Score 104.5; DB 3; Length 1043;
Best Local Similarity 20.5%; Pred. No. 0.86; 203; Indels 227; Gaps 35;
Matches 125; Conservative 54; Mismatches 54
Qy 5 SNDANPSDGTANLVPEVNN-----EVMALPEVVGAAIAAPVAGQQNVDPWIRNN 55
Db 37 SND-EPFGSQAGQIADTSNLFPPVQTHKSTGLPIDPMWGLPF-DPKSG--NLVHPYTN-- 90
Qy 56 FVQAFPG-----EFTVSPRNA---PGEILWSAPLGP---DLNPLYSHL-----AR 94
Db 91 --QTMWSGLSVSYLAAKNLTVDTDTYGLFDITLTGYPLDPVSLIPFPNPGELFDPIDSE 148
Qy 95 MYNGYAGGFEVQVILAGNAFTAGKIIFAAVPPNF-----PTEG-----LSPSQV 138
Db 149 INMGTIAGI-VSGISASELLSQKALIDPATNMVWVGEFGLLNPATGMIPGFLGPSEQ 207
Qy 139 TMFPHIIVDVRQLEP-----VLIPLDVRNNFYHYNQSNDSITKILAMLYPLR 187
Db 208 TQFSPEIEDGGIIPPEVAANADKFKLSIP-PSVPESIPKDKIDSISEL---MYDIES 263
Qy 188 ANNAGDDVTVCRLTRSP-----DPDFILVPTVESRT-KPFTVPILTVEEMS NR 241
Db 264 GRLLG-----QVSKRPGPSIAGDLNPIMKTTQTDSVTGKPID----- 302

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 34.4562 Seconds
(without alignments)
4368.312 Million cell updates/sec

Title: US-09-926-799-5
Perfect score: 2886
Sequence: 1 MMSANDANPSDGSSTANLVP.....VNQFTYLPAMGNGAGRRRAL 539

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/ECT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/ECT05_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1161.5	40.2	530	14	US-10-314-739-3
2	322	11.2	669	14	US-10-209-507-2
3	300.5	10.4	547	12	US-10-670-695-4
4	300.5	10.4	671	12	US-10-670-695-2
5	297.5	10.3	668	14	US-10-209-507-4
6	135	4.7	6310	12	US-10-282-122A-67793
7	126.5	4.4	1147	15	US-10-327-481A-38
8	122	4.2	26926	9	US-09-759-508B-2
9	116.5	4.0	913	15	US-10-369-493-2294
10	116	4.0	5701	9	US-09-864-761-37319
11	116	4.0	5701	14	US-10-029-386-32438
12	112.5	3.9	710	15	US-10-094-749-2315
13	112	3.9	1344	9	US-09-738-628-6888
14	111.5	3.9	776	9	US-09-765-272-86
15	111.5	3.9	2233	10	US-09-769-787-2

16	109	3.8	607	14	US-10-320-769-4	Sequence 4, Appli
17	109	3.8	782	9	US-09-908-193-47	Sequence 47, Appli
18	109	3.8	782	15	US-10-190-115-39	Sequence 39, Appli
19	109	3.8	782	15	US-10-369-072-39	Sequence 39, Appli
20	109	3.8	3930	12	US-10-282-122A-46817	Sequence 46817, A
21	108.5	3.8	1325	12	US-10-282-122A-44953	Sequence 44953, A
22	108.5	3.8	1767	12	US-10-087-684-40	Sequence 40, Appli
23	108.5	3.8	1767	12	US-10-087-684-41	Sequence 41, Appli
24	108.5	3.8	1767	12	US-10-218-779-40	Sequence 40, Appli
25	108.5	3.8	1767	12	US-10-218-779-41	Sequence 41, Appli
26	107	3.7	1323	14	US-10-156-761-10065	Sequence 10065, A
27	107	3.7	2481	12	US-10-282-122A-43762	Sequence 43762, A
28	106.5	3.7	866	12	US-10-282-122A-70192	Sequence 70192, A
29	106.5	3.7	1119	12	US-10-425-114-53646	Sequence 53646, A
30	106.5	3.7	1143	12	US-10-424-599-237981	Sequence 237981, A
31	106.5	3.7	1439	12	US-10-282-122A-61196	Sequence 61196, A
32	106.5	3.7	2669	12	US-10-016-248-4	Sequence 4, Appli
33	106.5	3.7	3104	12	US-10-016-248-2	Sequence 2, Appli
34	106	3.7	796	9	US-09-801-368-30	Sequence 30, Appli
35	106	3.7	1119	14	US-10-245-802-12	Sequence 12, Appli
36	105.5	3.7	5795	9	US-09-815-242-12610	Sequence 12610, A
37	105.5	3.7	10421	12	US-10-282-122A-61631	Sequence 61631, A
38	105	3.6	925	15	US-10-108-260A-3889	Sequence 3889, Ap
39	105	3.6	1313	12	US-10-282-122A-76863	Sequence 76863, A
40	104.5	3.6	1047	14	US-10-245-802-22	Sequence 22, Appli
41	104.5	3.6	1082	15	US-10-369-493-17890	Sequence 17890, A
42	103.5	3.6	685	15	US-10-104-047-2916	Sequence 2916, Ap
43	102.5	3.6	1794	10	US-09-965-738-299	Sequence 299, App
44	102.5	3.6	1799	10	US-09-965-738-149	Sequence 149, App
45	102.5	3.6	1821	10	US-09-965-738-82	Sequence 82, Appli

ALIGNMENTS

RESULT 1
US-10-314-739-3
; Sequence 3, Application US/10314739
; Publication No. US20030129588A1
; GENERAL INFORMATION:
; APPLICANT: Estes, Mary K
; Jiang, Xi
; Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; Characterize No. US20030129588A1walk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Ave., N.W.
; CITY: Washington, D.C.
; STATE: <Unknown>
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/10/314,739
; APPLICATION NUMBER: US/10/314,739
; FILING DATE: 09-Dec-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,049
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311,023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX: <Unknown>

; INFORMATION FOR SEQ ID NO: 3:		; SOFTWARE: PatentIn version 3.0	
; SEQUENCE CHARACTERISTICS:		; SEQ ID NO 2	
; LENGTH: 530 amino acids		; LENGTH: 669	
; TYPE: amino acid		; TYPE: PRT	
; TOPOLOGY: linear		; ORGANISM: Feline calicivirus	
; MOLECULE TYPE: protein		US-10-209-507-2	
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:			
US-10-314-739-3			
Query Match		Query Match	
Best Local Similarity 40.2%; Score 1161.5; DB 14; Length 530;		Best Local Similarity 25.4%; Pred. No. 1.5e-20;	
Matches 250; Conservative 83; Mismatches 181; Indels 37; Gaps 14;		Matches 125; Conservative 69; Mismatches 158; Indels 140; Gaps 22;	
QY	1 MMWASNDANPS-DGST--ANLYPEVN-NEVMALEPVVGAATAAPVAGQONVIDPWIRNPF 56	QY	11 SDGSTANLPEVNVEY--MALEPVVGAATAAPVAGQONVIDPWIRNPFVQAPGGEFTV-- 66
DB	1 MMWASKDATSSVDGASGAGQLPEVNASDPLAMDPLVAGSSTAVATAGQNPIDPWIIINF 60	DB	125 ADGDSSTITPEQGTLLGVVIAEPFSAQMAATAAATGKSDSEW--ESP-----FSFHT 175
QY	57 VQAPGGEFTVSPNAPGELTWSAPLGPDLNPNYLSHLARMYNGYAGGFEVQVILAGNAFTA 116	QY	67 -----SPRNPAGELTWSAPLGPDLNPNYLSHLARMYNGYAGGFEVQVILAGNAFTA 122
DB	61 VQAPQGEFTISPNNTPGDVLFDLSLGFHLNPFLLHLSQYNGWGVNMRVIRIMLAGNAFTA 120	DB	176 SVNWSSETQKLLFKQSLGLPLNPLYLHLSKLYVWASGVSVRFISGSGVFGKLAAL 235
QY	117 GKIIIFAAVPPNPETEGLSPSQVTMPFPHIIVDROLEPVLIPLDVRRNNFYHYNQSDSTI 176	QY	123 AVPPNPTEGLSPSQVT---MPFPHIIVDROLEPVLIPLDVRRNNFYHYNQSDSTIKLI 179
DB	121 GKLIIVSCIPPGFSGHNTIAQATLFPHVIAVRLDPIEVPLEDVRNVLFNHNRNQOTM 180	DB	236 VVPP-----GVDPVQSTSLMLQYPHVLFDAQVPEFVIFSPDLKSTLHLMSDTHL--SLV 289
QY	177 KLIALMYPLRANAGDDVFTVSCRVLTRPSDPDFILVPPVPSRTKPTFVPILTVEE 236	QY	180 AMLYTPLRANAGDDVFTVSC--RVLTRPSDPDFILVPP-----TVES----- 222
DB	181 RLVCMLYPLRLTGGTGDGDSFVAVRVTCPSPDFNLFELVPPVTEOKTRPFTLPLNPLSS 240	DB	290 IMVYNDL--INPYANDSNSSCIVTVETKPGDPKFKHLKPPGSMLTGHSIPSDLI 348
QY	237 MNSRRPPILEKLYTGPSSAFVQPNQGRCTTGVLLGTTQLSAVNICTFRGVDTHIAGS 296	QY	223 -----RTKPTFTVPILTVEEMSRSRFP 243
DB	241 LNSRAPLPISSGISPDNVQSVQFQNGRCITLDGLRLGITPVSLSHVAKIRGT-----S 294	DB	349 LWIGNRYSDITDPIRPFVFOANRHFDFNQETAGMSTPRFRPITI-----TISESNGS--- 402
QY	297 HDYTMNLASQWNSYNDTEBIPAPLGPDPVGVKIQKMLTQTTREDGSTRAHKATVSTGVS 356	QY	244 IPLEKLVTPGSSAPVQV--PQNGRCITDGLVLTQLSAVNICFERGDVTHIAGSHDYTM 301
DB	295 NGTVINLTDLGTPFHFEG-PAPIGPPD-LGGCDWHINMT--QFGHSSQOTQVDVDTTPD 350	DB	403 ----KLGTGVATDVIIVPGIPDGPDPDTTIGBEL-----TPAGDVIITNGSGN--- 444
QY	357 HFTPKLGSVQYTTDNTNDFOTGQNTKFTPVGVQ--DGNHNEPO--QWVLPNYSGRTG 412	QY	302 NLASQWNSYNDTEBIPAPLGPDPVGVK-IQGMLTQ-----TTREDGSTR 346
DB	351 TFPVHLGSIQ-----ANGISGNY-----GVLSWISPPSPHSGSQVDLWKIPNYGSSIT 400	DB	445 DIATAN--AYDSADVI---TNTNFRGMWICGALQRAWGDKKISSTAFITTAKEGNTLK 499
QY	413 HNVHLAPAVPTFPGEQLLFRSTMPGCGYPNNMLDCLLPQEWVHFQCOEAPAQSDVA 472	QY	347 HKATVSTGSHFTPKLGSVQYTTDNTNDFOTGQNT-----KFTPVGVIQDGNHNEPQOW 402
DB	401 EATHLAPSVYPPGGEVLVFMSSKMPGPGAY--NLPCLLPQEVISHLASEQAPTVEGA 457	DB	500 PSNTIDMTKI-----AVYQDTHVGRDVQTSDDTLAILGYTGIGEQAIQSNRDSVVRI 552
QY	473 LLRFVNPDTGRVLPCKLHKSQYTV-----AHTGPHDLVTPNGYFRFDSWVNOFTLAP 528	QY	403 VLPNYSGRTGN 414
DB	458 LLHYVDPDTGRNLGEFKAYPDGFLTCVPNGASSGPQOL--PINGVVFVSVWSRFPQLXP 515	DB	553 MLPETGARGGNH 564
QY	529 MNGAGRRERL 539		
DB	516 VGTASSARGRL 526		
RESULT 2		RESULT 3	
US-10-209-507-2		US-10-670-695-4	
; Sequence 2, Application US/10209507		; Sequence 4, Application US/10670695	
; Publication No. US20030109033A1		; Publication No. US20040058316A1	
; GENERAL INFORMATION:		; GENERAL INFORMATION:	
; APPLICANT: Audonnet, et al.		; APPLICANT: Jensen, Wayne A.	
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT		; APPLICANT: Lappin, Michael R.	
; FILE REFERENCE: 454313-3151.2		; APPLICANT: Rosen, David K.	
; CURRENT APPLICATION NUMBER: US/10/209,507		; APPLICANT: Andrews, Janet S.	
; CURRENT FILING DATE: 2002-11-12		; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE	
; PRIOR FILING DATE: 2000-07-14		; FILE REFERENCE: DI-9-1	
; PRIOR APPLICATION NUMBER: 60/193,332		; CURRENT APPLICATION NUMBER: US/10/670,695	
; PRIOR FILING DATE: 2000-03-30		; CURRENT FILING DATE: 2003-09-25	
; PRIOR APPLICATION NUMBER: France 00 01761		; PRIOR APPLICATION NUMBER: 09/521,738	
; PRIOR FILING DATE: 2000-02-11		; PRIOR FILING DATE: 2000-03-09	
; PRIOR APPLICATION NUMBER: France 99 09421		; NUMBER OF SEQ ID NOS: 36	
; PRIOR FILING DATE: 1999-07-16		; SOFTWARE: PatentIn ver. 2.1	
; NUMBER OF SEQ ID NOS: 26		; SEQ ID NO 4	
		; LENGTH: 547	
		; TYPE: PRT	
		; ORGANISM: Feline calicivirus	
		US-10-670-695-4	
Query Match		Query Match	
		10.4%; Score 300.5; DB 12; Length 547;	

Best Local Similarity 25.3%; Pred. No. 1.1e-18;
Matches 141; Conservative 80; Mismatches 238; Indels 99; Gaps 28;

QY 12 DGSIT--APEQGTWGGVIAEBSAQMSAADMATKGSVDSEW-----EAPFSFHTSVNM 178
DB 12 DGSIT--APEQGTWGGVIAEBSAQMSAADMATKGSVDSEW-----EAPFSFHTSVNM 178
QY 67 SPRNAPGELWSAPGLDPLNLYSHARMYAGGFEVQVILAGNAFTAGKIIFAAVPP 126
DB 67 SPRNAPGELWSAPGLDPLNLYSHARMYAGGFEVQVILAGNAFTAGKIIFAAVPP 126
QY 55 STSETQKILFKQSLGPLNLYSHARMYAGGFEVQVILAGNAFTAGKIIFAAVPP 238
DB 55 STSETQKILFKQSLGPLNLYSHARMYAGGFEVQVILAGNAFTAGKIIFAAVPP 238
QY 127 NPPTGLSPSQVT---MPPHIIVDVRQLEPVLPLPDVNNPYPHYNQSDSTIKILAMLY 183
DB 127 NPPTGLSPSQVT---MPPHIIVDVRQLEPVLPLPDVNNPYPHYNQSDSTIKILAMLY 183
QY 115 -----GVDPVQSTMLQYPHVLFDAQVPEVIFCLPDLRSTLYHLMSDITTT-SLVIMVY 292
DB 115 -----GVDPVQSTMLQYPHVLFDAQVPEVIFCLPDLRSTLYHLMSDITTT-SLVIMVY 292
QY 184 TPLRANNAGDDVFTVSC--RVLTSPSPDPDFELVPPTVESRTKPTVPILTVEMSNR 241
DB 184 TPLRANNAGDDVFTVSC--RVLTSPSPDPDFELVPPTVESRTKPTVPILTVEMSNR 241
QY 169 NDL--INPYANDANSSGCIIVTETKPGDPKFHLLKPP--GSMITHGSPSDLPKTS-- 347
DB 169 NDL--INPYANDANSSGCIIVTETKPGDPKFHLLKPP--GSMITHGSPSDLPKTS-- 347
QY 242 FPIPLEKLYTGSSAFVQVP---QNGR-----CTTDGVLGTTQLSAVNICTFRGDVTHI 293
DB 242 FPIPLEKLYTGSSAFVQVP---QNGR-----CTTDGVLGTTQLSAVNICTFRGDVTHI 293
QY 224 --LWIGNRYWSDITDFVIRPFVFOANRHFDFNQETAGWSTPRFRPISVTITEQNGAKLGI 405
DB 224 --LWIGNRYWSDITDFVIRPFVFOANRHFDFNQETAGWSTPRFRPISVTITEQNGAKLGI 405
QY 294 AGSHDYTNLASQNSNYD--PTEBIPAPLGTDFVKGIOGMLTQTTREDGSTRA--HKAT 350
DB 294 AGSHDYTNLASQNSNYD--PTEBIPAPLGTDFVKGIOGMLTQTTREDGSTRA--HKAT 350
QY 282 GVATDYIVPGIPDGWPDITIPGELIPA-----GDYA-----ITNGTGNDITATGYDTAD 455
DB 282 GVATDYIVPGIPDGWPDITIPGELIPA-----GDYA-----ITNGTGNDITATGYDTAD 455
QY 351 VSTGSHVF-----TPKLGSVQYTT-----DTNNDPQTQNTKFTVPVGIQ 390
DB 351 VSTGSHVF-----TPKLGSVQYTT-----DTNNDPQTQNTKFTVPVGIQ 390
QY 332 IIKNTNFRGMVYICGSLQRAWGDKKISNTAFITATLDGDNKINPCNTIDQSKIIVVQ 437
DB 332 IIKNTNFRGMVYICGSLQRAWGDKKISNTAFITATLDGDNKINPCNTIDQSKIIVVQ 437
QY 391 D---GNHQNHPQWVLPNYSG-----RTGHNHVLAPAVAPTPP-----GEQLLFFRSTM 437
DB 391 D---GNHQNHPQWVLPNYSG-----RTGHNHVLAPAVAPTPP-----GEQLLFFRSTM 437
QY 392 DNHVGKKAQTSDDTLALLGYTGIGEQAGSDDRVVRIS--TLPETGARGGNHPIFYKNSI 574
DB 392 DNHVGKKAQTSDDTLALLGYTGIGEQAGSDDRVVRIS--TLPETGARGGNHPIFYKNSI 574
QY 438 PCSGYPNNMLDCLLPQEWVQHFQEEA-----PAQSDVALLRFVNPDTGRVLFCKLH 491
DB 438 PCSGYPNNMLDCLLPQEWVQHFQEEA-----PAQSDVALLRFVNPDTGRVLFCKLH 491
QY 451 K--LGVIRSIDVFNQ--ILHRSQSLNHYLLPPDS--FAVYRII--DSNGSWPFDIGID 627
DB 451 K--LGVIRSIDVFNQ--ILHRSQSLNHYLLPPDS--FAVYRII--DSNGSWPFDIGID 627
QY 492 KSGYVTVVHTGPHDLVIP 509
DB 492 KSGYVTVVHTGPHDLVIP 509
QY 504 SDGFSFVGVSGFGKLEFP 521
DB 504 SDGFSFVGVSGFGKLEFP 521

RESULT 4
US-10-670-695-2
; Sequence 2, Application US/10670695
; Publication No. US2004005816A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; PRIOR FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 671
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-670-695-2

Query Match 10.4%; Score 300.5; DB 12; Length 671;
Best Local Similarity 25.3%; Pred. No. 1.5e-18;
Matches 141; Conservative 80; Mismatches 238; Indels 99; Gaps 28;

QY 12 DGSIT--APEQGTWGGVIAEBSAQMSAADMATKGSVDSEW-----EAPFSFHTSVNM 178

DB 127 DGSIT--APEQGTWGGVIAEBSAQMSAADMATKGSVDSEW-----EAPFSFHTSVNM 178
QY 67 SPRNAPGELWSAPGLDPLNLYSHARMYAGGFEVQVILAGNAFTAGKIIFAAVPP 126
DB 179 STSETQKILFKQSLGPLNLYSHARMYAGGFEVQVILAGNAFTAGKIIFAAVPP 238
QY 127 NPPTGLSPSQVT---MPPHIIVDVRQLEPVLPLPDVNNPYPHYNQSDSTIKILAMLY 183
DB 239 -----GVDPVQSTMLQYPHVLFDAQVPEVIFCLPDLRSTLYHLMSDITTT-SLVIMVY 292
QY 184 TPLRANNAGDDVFTVSC--RVLTSPSPDPDFELVPPTVESRTKPTVPILTVEMSNR 241
DB 293 NDL--INPYANDANSSGCIIVTETKPGDPKFHLLKPP--GSMITHGSPSDLPKTS-- 347
QY 242 FPIPLEKLYTGSSAFVQVP---QNGR-----CTTDGVLGTTQLSAVNICTFRGDVTHI 293
DB 348 --LWIGNRYWSDITDFVIRPFVFOANRHFDFNQETAGWSTPRFRPISVTITEQNGAKLGI 405
QY 294 AGSHDYTNLASQNSNYD--PTEBIPAPLGTDFVKGIOGMLTQTTREDGSTRA--HKAT 350
DB 406 GVATDYIVPGIPDGWPDITIPGELIPA-----GDYA-----ITNGTGNDITATGYDTAD 455
QY 351 VSTGSHVF-----TPKLGSVQYTT-----DTNNDPQTQNTKFTVPVGIQ 390
DB 456 IIKNTNFRGMVYICGSLQRAWGDKKISNTAFITATLDGDNKINPCNTIDQSKIIVVQ 515
QY 391 D---GNHQNHPQWVLPNYSG-----RTGHNHVLAPAVAPTPP-----GEQLLFFRSTM 437
DB 516 DNHVGKKAQTSDDTLALLGYTGIGEQAGSDDRVVRIS--TLPETGARGGNHPIFYKNSI 574
QY 438 PCSGYPNNMLDCLLPQEWVQHFQEEA-----PAQSDVALLRFVNPDTGRVLFCKLH 491
DB 575 K--LGVIRSIDVFNQ--ILHRSQSLNHYLLPPDS--FAVYRII--DSNGSWPFDIGID 627
QY 492 KSGYVTVVHTGPHDLVIP 509
DB 628 SDGFSFVGVSGFGKLEFP 645

RESULT 5
US-10-209-507-4
; Sequence 4, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-4

Query Match 10.3%; Score 297.5; DB 14; Length 668;
Best Local Similarity 26.8%; Pred. No. 2.8e-18;
Matches 110; Conservative 52; Mismatches 147; Indels 101; Gaps 20;

QY 12 DGSIT--APEQGTWGGVIAEBSAQMSAADMATKGSVDSEW-----EAPFSFHTSVNM 178
DB 127 DGSIT--APEQGTWGGVIAEBSAQMSAADMATKGSVDSEW-----EAPFSFHTSVNM 178

ORGANISM: Pseudomonas putida
US-10-282-122A-67793

Query Match 4.7%; Score 135; DB 12; Length 6310;
Best Local Similarity 23.7%; Pred. No. 0.11;
Matches 108; Conservative 50; Mismatches 189; Indels 108; Gaps 24;

QY 7 DANPS-----DGTANLYPEVNEVMALEPVVGAIAAPVAG---QQNVDPWIRNPFVQ 58
DB 3575 NASPSATFTADDSTAPAA-----AGLTTPGSAIQSGEAGSTVEVKLADGTLVGVTVV 3630

QY 59 AGGEFTV--SPRNAPGEIL-----WSAP---LGPDLNP-----YLSHLARMY 96
DB 3631 PAGGSTVPLSPAQLDQALNVLTLDAAAGNISQPSQIPAPDITPPALFTDVAVSSDGTAV 3690

QY 97 NGYAGFEVQVILAGNAFTAGKIIIPAAVPPN-----FPTGLSPSOVTMPFPHIIVDV-- 148
DB 3691 TGNAPGASSVTVSDG-----AGNVITVAVNPDSGFSVPLDTPONNGQTVTV--VVTDAAG 3743

QY 149 RQLEPVLILPLPDVRNMFYHN---QSNSTIKILIAMLYTFLRANNAGDDVFTVSCRVLTR 205
DB 3744 NDSAPVSVTAPDTNPEPATGLTVSPDGSTVCGTAEPGSTVEVRNPDDTV-----RGTVT 3798

QY 206 PSPDFDFILVPTVESRTKPTVPIITVEEMSNRFPPILEKLYTGPSAFVWOPQNGR 265
DB 3799 AGPDGTFVIVAPPLASGE---TVDVVVIDPAGNESPEIPL---TGPTGTEVATPSALA 3851

QY 266 CTTDGVLLGTTLQSLAVNICTFRGDVTTHIAGS---HDYTMNLASQN-----WSN 310
DB 3852 ISVDGFL--TGQGTGSLITVTSGGTTLGSAVTVSGDTPRVFFQNAQLNAQILQVSAKAT 3910

QY 311 YDPTTEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAH-KATVST-----GSVHFTPKLGSV 365
DB 3911 VDGQPSVPAIIVAND-----TTAPDAPTQVNLNATGTLTGQGEVGTATVRVTDL 3959

QY 366 Q-----YTTDTNNDFTQNTKFTPVGVIOQDGN 394
DB 3960 QGTLTGATVDSNGLF-----SVSFSP--AVANGQN 3988

RESULT 7
US-10-327-481A-38
; Sequence 38, Application US/10327481A
; Publication No. US20040001864A1
; GENERAL INFORMATION:
; APPLICANT: King, Andrew M.Q.
; APPLICANT: Burman, Alison J.
; APPLICANT: Audonnet, Jean-Christophe F.
; APPLICANT: Lombard, Michel F.A.
; TITLE OF INVENTION: Vaccine against Foot-and-Mouth Disease
; FILE REFERENCE: 454313-3178
; CURRENT APPLICATION NUMBER: US/10/327,481A
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: PCT/FR01/02042
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: FR 0008437
; PRIOR FILING DATE: 2000-06-29
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 1147
; TYPE: PR1
; ORGANISM: Foot-and-mouth disease virus
US-10-327-481A-38

Query Match 4.4%; Score 126.5; DB 15; Length 1147;
Best Local Similarity 22.2%; Pred. No. 0.048;
Matches 78; Conservative 41; Mismatches 144; Indels 89; Gaps 15;

QY 72 PGETLSAPGLPDLNPLVLSHARMYNGYAGGFEVQVILAGNAFTAGKIIIPAAVPPNPFTE 131
DB 160 PFGYLTKELEPDTDHHGVGHVDSYAMRMGMDEVSAGVNOFNGGCLLIVAMVPEWKAFD 219

US-10-282-122A-67793

QY 65 TVSPRNAPGEILMSAPLGDPLNPLVLSHARMYNGYAGGFEVQVILAGNAFTAGKIIIPAAV 124
DB 177 NWMTSETQKILFKQSLGFLNPLVLSHARMYNGYAGGFEVQVILAGNAFTAGKIIIPAAV 236

QY 125 PNFTEGLSPSOVT---MFPHIIVDVRLQLEPVLILPLPDVRNMFYHNQSNDSSTIKILIAM 181
DB 237 PP-----GIDPVQSTSMLOQPHVLFDAQVPEVFTIIDLNSLYLHMSDDTDT-SLVIM 290

QY 182 LYTPLRANNAGDDVFTVSC--RVLTRPSPDFDFILVPTVESRTKPTVPIITVEEMSN 239
DB 291 IYNDL-INPYANDSNSGCIIVTETKPGDFKPHLLKPPG-----SMLTHGSIPS 339

QY 240 SRPIPLEKLYTGPSAFVQONGRCITDGVLLGTTLQSLAVNICTFRGDVTTHIAGSHDY 299
DB 340 DLIP-----KSSSLWIGNRHMSDITDFV-----IKPVFQ-----ANRH-F 374

QY 300 TNNLASQNSNVDPTTEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGCV--- 356
DB 375 DFNQETAGNS-----TPRF-----RPITITVSEKGSKL-GIGVATDSIVPG 415

QY 357 -----HFTPKLGSVQVTTDTNNDFTQNTKFTPVGVIOQDGN 394
DB 416 IPDGWPDITTEPKLTP-AGDYAITNGGNDITTAAD--YDGASILKNNTN 462

RESULT 6
US-10-282-122A-67793
; Sequence 67793, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67793
; LENGTH: 6310
; TYPE: PR1

QY 132 GLSPQVTFPHIIVDVROLEPVLPLDVRNFF-HYNQSDSTIKLIAMLYTPIRANN 190
DB 220 TREKQLTLFPHQFISPRNTMAHTIYVYLGVRNRYDQYKHKPWT--LVVMVLSPLTVSN 277
QY 191 AGDDVFTVSCRVLTRPSDFDFILVPPFTV-----ESRTKPTVPILITVEEMSNSRFP 245
DB 278 TA-----APOIKYANIAPYVHVAGELPSKEGIFPVACADGGLVTTDP 323
QY 246 -----LEKLYTGPSSAFVVPQNGRCTTGDVLLGTQLSAYNIC--TF----- 286
DB 324 KTADPVYGVKNPPKTYN-----PGRF-----TNLLDVAEACPTFLRPDDGKPYVVT 370
QY 287 RGDVTHIAGSHDYTNLASQNSVNDPTEIEPAPLGTDPFVGKIQOGLTQTTREDGSTR 346
DB 371 RADDTRLLAKFD--VSLAKHMSN-----TYLSGIAOYYTOYSGTINL 411
QY 347 H-KATVSTGSGS-----VHFTPKLGSQYTTDT-----NNDFTQGTQNTKET 384
DB 412 HFMFTGSTDSKARYWVAYIPP--GVETPPDTPEEAHCHIAEWDGLNSKFT 461

RESULT 8
US-09-759-508B-2
; Sequence 2, Application US/09759508B
; Publication No. US20020182599A1
; GENERAL INFORMATION:
; APPLICANT: Fishman, Mark C.
; TITLE OF INVENTION: Methods for Diagnosing and Treating Heart Disease
; FILE REFERENCE: 00786/381002
; CURRENT APPLICATION NUMBER: US/09/759,508B
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/175,787
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 26926
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-508B-2

Query Match 4.2%; Score 122; DB 9; Length 26926;
Best Local Similarity 19.5%; Pred. No. 17;
Matches 118; Conservative 83; Mismatches 211; Indels 194; Gaps 30;

QY 63 EFTVSPRNPAG-----ETLM--SAPLGPDLNPLY-----SHLARMYGVAGGF 103
DB 14146 EPRVSAENRAGIPSPSEASDSVLMKDAAYPPGPPSNPHVDTTKKSASLWKGPHYDGL 14205
QY 104 EV-----QVILAGNAFTAGKIIFAAVPP-----NFTPEGLSPSQV---TMF 141
DB 14206 EITGVVVEHQVGDENWIKDTGTALRITQFVVPDLQTKYKNFRISAINDAGVGEPAVI 14265
QY 142 PHIIVDVROLEP-----VLIPL-----PDV-----RNNFYHYNOSN 172
DB 14266 PDVEIVEREMAPDFELDAELRRLTVVRAGLSIRIFVPKGRPAPEVTWKONILNKRAN 14325
QY 173 DSTIKLIAMLYTP-----LRANNAGDDVFTVSCRVLTRPSDFDFILVPP-- 218
DB 14326 INTESFTLLIIPENRVDYTGKVMTIENPAGKSGFVNVRVLDTPGP---VNLNRPTDI 14382
QY 219 TVESRTKPTVPILITVEEMSNSRFPITPLEKLYTGPSSAFVVPQNGRCTTGDVLLGTQL 278
DB 14383 TKDSVTLHWDPLID-----GGSRI-----TNVIVEKR--EATKRSYSTATIK- 14423
QY 279 SAVNICTFRGDVTHIAGSHDYTNLASQNSVNDPTEIEPAPLGTDPFVGK 329
DB 14424 --CHKTYK--VTGLSEGCYEFVRMAENYGEIPEPTETTEPVKASEAPSP---PDSLNI 14476
QY 330 IQOGLTQTT-----RDGSTRAHKATVSGSVHFTPKLGSQYTTDTN-----N 373
DB 14477 MD--ITKSTVSLAWPKPKHGGSKI-----TGYVIEAQKSGSDQWTHITTVKGLECVVR 14528

RESULT 9

US-10-369-493-2294
; Sequence 2294, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 2294

; LENGTH: 913

; TYPE: PRT

; ORGANISM: Schizosaccharomyces pombe

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(913)

; OTHER INFORMATION: unsure at all Xaa locations

US-10-369-493-2294

Query Match

4.0%; Score 116.5; DB 15; Length 913;

Best Local Similarity 22.4%; Pred. No. 0.29;

Matches 99; Conservative 48; Mismatches 190; Indels 105; Gaps 22;

QY 10 PSDGSTANLVPEV-----NNEVMALPEVVGAAIAAPVAGQQNVIDPWIRNNFVQAP-GGE 63

DB 490 PAVGSTTTSTVEVPLAGNVEVYPTPEPTMVSIVTPVAVGSTTSTVPEV-----VPLAGN 543

QY 64 FTVSPRNAPGELMSAPGLDNLNLYSLHARMYGVAGFEVQVILAGNAFTAGKIIFAA 123

DB 544 VEYVTEPETMVSIVTPVAVGSTNSTVPEVPL-----AGNVEV-----NALEPDTMVSIV 593

QY 124 VPPNPTEGLSPSQVTFPHIIVDVROLEP---VLIPLPDVRNFFHYHYNOSNDSTIKLIA 180

DB 594 PAVGSTTTSTVPEVPLAGN--VEVKLEPETMVSIVPAV-----GSTTSTVPEV- 642

QY 181 MLYTPLRANNAGDDVFTVSCRVLTRPSDFDFILVPPPTVESRTK--PFTVPILTVEEMS 238

DB 643 ---VPLAGN-----VEYVTEPETMVSIVPAVGSTTSTVPEVPLAGNVEV- 686

QY 239 NSRPFIPLEKLYTGPSSAFVVPQNGRCTTQGV-----LLGTTQLSASVNICTFRGDVT 293

DB 687 -----KYLEPETMVSIVPAVAVG-STTSTVPEVPLAGNVEVNALEPDTMVSIVPA 736

QY 294 AGSHDYTMNLASQNSVNDPTEIEPAPLGTDPFVGKIQOGLTQTTREDGSTRAHKATVST 353

DB 737 VGSTTSTV-----PEVPLAGNVEYVTEPETMVSIVTPVAVGST-----T 776

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2
OTHER INFORMATION: SWISSPROT HIT: P16419, EVALUE 2.00e-72
US-10-029-386-32438

Query Match 4.0%; Score 116; DB 14; Length 5701;
Best Local Similarity 19.2%; Pred. No. 5.4;
Matches 111; Conservative 81; Mismatches 203; Indels 182; Gaps 28;

QY 80 PLGPDLPYL-----SHLARMYNGYAGGFV-----QVILAGNAFTAGKII 120
DB 1 PGPSPNPHVTDTTKASLANGKHYDGLSITGVVHEHQKGVDEAWIKOTGTALRIT 60
QY 121 FAAVPP-----NFTPEGLSPSQV---TMFPHIIVDVRLQLEP----- 153
DB 61 QFVVPDLQTKYKNFRISAINDAGVCEPAVIDVEIVEREMAPFELDAELRTLVRAG 120
QY 154 -----VLIPL-----PDV-----RNFVHYNOSNSTIKLIAMLYTP-----LRAN 189
DB 121 LGSIRFVPIKGRPAPEVTTKONILNKRANIENTESFTLLIIPECNRYDTGKFMVTIEN 180
QY 190 NAGDDVFTVSCRVLRPSDFDFELVPP--TVESRTKPTFTVPILTVEMSNSRPIPLE 247
DB 181 PAGKSGFVNVRLDTGPG---VLNLRPTDITKDSVTLHWDLPID-----GGSRI----- 228
QY 248 KLYTGPSSAFVQVQNGRCTTGVLLGTQLSAVNICTPRGDVTIAGSHDVTMNLSON 307
DB 229 -----TNYIIVEKR--EATRSKSYATYK---CHKTYK--VTGLSECEYFPFVMAEN 274
QY 308 -MSNYDPT-----ELPAPLGTDFVKGIOGMLTQTT-----REDGSTRAHKAT 350
DB 275 EYGEIETETTEPVKASEAPSP---PDSLNIIMD--ITKSTVSLAWPKPKHGGSKI----- 325
QY 351 VSTGSHFTPKLGSVQYTTDTN-----NDFQTGNTKFTFVGVIOGNNHNE----- 398
DB 326 --TGVIEAQRKSGDOWHTITTVKGLCEVVRNLTEGEYTFQMAVNSAGRSAPRESRPV 383
QY 399 ---PQOWLPNYS-----GRTHNVHLAPV-----APTF---PGEQLLFFRSTWPG 439
DB 384 IYKEQTMPELDRLGIYQKLVITAKAGDNIKVEIPVLGRPKFTVWKKGQDIL----- 435
QY 440 CSGYPNMNLDCLLPQEWQHFQCEAAPQSDVALLRFVNPDTGRVLFECKLHKSGYVTA 499
DB 436 -----KQQRVNFETTATSTILNINECVRSDSGPYPLTARNIVGEVDVI 480
QY 500 HTGPHDLVLPNGYFRFDSWVNOFTLA---PMNGGAG 534
DB 481 TIQVHDIPGPTGPIKDFDEVSSDFVTFSDPPENDGG 517

RESULT 12

US-10-094-749-2315
Sequence 2315, Application US/10094749
Publication No. US20030219741A1

GENERAL INFORMATION:

APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOKUKI

APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2315
LENGTH: 710
TYPE: PRT
ORGANISM: Homo sapiens
US-10-094-749-2315

Query Match 3.9%; Score 112.5; DB 15; Length 710;
Best Local Similarity 21.3%; Pred. No. 0.45;
Matches 97; Conservative 45; Mismatches 163; Indels 151; Gaps 19;

QY 3 MASNDANPSDGSSTANLVPENNEV--MALEPVVGAIAAPVA---GQONVIDPWIRNFVQ 58
DB 123 MPSRFIAAPPSTLSPKVPQGVTVTMESSIPQASAIPTATISGQGG----- 170
QY 59 APGGFTVSPRNAPGEILWSAPLGPDLNPLYLHARMYNGVAGGFEVQVILAGNAFTAGK 118
DB 171 HPSNLHHIMTNTVQMSIIRSNAPEGPLHIGASHLPR---GAAA----- 210
QY 119 IFAAVPNPTEGLSPQSVTMFPHIIVDVRLQLEVLPLPDPVRNPFVHYNOSNSTIKL 178
DB 211 ---AAV-----MSSSKVTT--VLRPTSQLPNAATAQPAVOHIIHQPIQSR----- 250
QY 179 IAMLVTPLRANAGDDVFTVSCRVLTRSPDFLFLVPPTV---ESRTKPTVPILTVE 235
DB 251 -----PPVTTGNA-----IPPAVAVTVSATRAQSPVITTTA 281
QY 236 EMSNRFFIPEKLYTGPSSAFVQV--PQNGRCTTGVLLGTQLSAVNICTPRGDVTIHA 294
DB 282 AHATDSALRPTLSIOHPFSAAISIORPAQSRDVTITL-----PSHPA 325
QY 295 GSHDYTMNLSONWSNYDTEIPEIPLCTPDPFVGKIOGMLTQTTREDGSTRAHKATVSTG 354
DB 326 -----LGTPKQQLHTMAKTIPTGTGTGTVAAATVAPILATNTIPSTATT-----AG 369
QY 355 SVHFT--PKLGSVQYTTDNTNDFQTGNTKFTFVGVIOGNNHNEPQOWLPNYSGRGT 412
DB 370 SVSHQTAPTSTIVTVPFSSHSHATAVTTSNIPVAKV-----VPOO----- 410
QY 413 HNVHLAPAVAPTFPGEQLLFFRSTWPGSGY---PN 445
DB 411 -ITHTSPRIQPDYPAE-----RSSLIPISGHRASPN 440

RESULT 13

US-09-738-626-6888
Sequence 6888, Application US/09738626
Publication No. US20020197605A1

GENERAL INFORMATION:

APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOHO, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OKAZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626

Wed Jun 2 09:13:30 2004

ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLY APPLICATION NUMBER: US/09/765,272
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLY APPLICATION NUMBER: 08/961,083
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:
LENGTH: 776 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 86:
US-09-765-272-86

Query Match 3.9%; Score 111.5; DB 9; Length 776;
Best Local Similarity 19.9%; Pred. No. 0.65;
Matches 111; Conservative 63; Mismatches 235; Indels 149; Gaps 25;

QY 5 SNDANPSDGSSTANLVPEVNEVMALEPVVGAATAAPVAGQONVIDPWIRNPFVQAGGEF 64
DB 62 ASDSNPSD-----PVS-----NVNDKLIISYN--NQPANRW 89
QY 65 T-----VSPRNPAGSILMSAPLGPDLNPLYSLHARMYAGGFEVQVILAGNAFTAKII 120
DB 90 TNWARTNPEASVGVLFDSGI-----LS--KESVDNLSVGFHED--HGVGVKPSYVI 137
QY 121 PAAVPPNPPTGSLPSQVTMPFPHIIVDVRLQLEPVL-----IPLPD 160
DB 138 EYTVGKTVPTAPKNPSFVGNEDHVFNDSEANWKPVTNLKAPALQKAGEMNHFSFDKVTYA 197
QY 161 VRNNFYHYNQSDSTIKLIAMLYTPLRANNAGDDVFTVSCRVLTRPSPDFDFIFLVPPTV 220
DB 198 VRIRMVADKNKRGTSITEVQIFAKOVAAKOGQOTRIQVDGKOLANFNPDLDYYL--ESV 255
QY 221 ESRTPKFTVPIITVEEMSNRPPIPLEKLYTGPSSAFVVPQNG-----RCTDGLV 272
DB 256 DGR-----VPAVTA-SVNNGLATVVPVREGEPRVIAKAENGDI LGELRHFTKDKSL 309
QY 273 LGTTQLSAYNICTFR--GDV-----THIAGSHDYTNLNASQWNSVNDPTEEPAPLG 322
DB 310 LSHKPVAAVKQARLLQVGOALELPTKVPVYFTGKDGYTEKDLTVEW-----EVPANL 363
QY 323 TPDFVGKIQGMLTQTTREDGSTRAHKATVSTGSHFTPKLGSV-----QYTTDNDPQT 377
DB 364 TK-----AGQFTVRGRVGLGSNLVAEIT-----VRVTDKLGTELSDNPNYDENSNAQAFAS 412
QY 378 GQN-----TKFTPVGVIQDGNHNEPQQWLVNPNYSGRTCHNVHLAPAVAPTPEGQLLFF 433
DB 413 ATNDIDKNSHDRVVDYLDNDGDHSEN--RRWT--NWSPTSSNPEVSAGV-----IF 458
QY 434 RSTWPGCSGYFNNDLCLLPQEWVQHFCEAAAPQSDVALLRFVNPDTGRVLFECKLHKS 493

QY 1 MKMASNDANPSDGS--TAN-LVPE-----VNEVNEVMALEPVVGAATAAPVAGQONVIDPW 52
DB 158 VEITNTAPIDFSDYTLNLYPQDEFTTNEAVAAEP--GDVLIQP--GKSLVF--WI 211
QY 53 RNFVQAGGEFTVSPRNPAGSILMSAPLGPDLNPLYSLHARMYAGGFEVQVILAGN 112
DB 212 KN-----GPNDEATAADFNA--EYGTNLEAGKDLVEISS--GGMANGTARQMGIQT-NTGH 262
QY 113 AFTAGKIIFAAVPPNPPTGSLPSQVTMPFPHIIVDVRLQ-LEPVLI-----PLPDVRNPFVH 167
DB 263 IVNRGFYNMAGSDVKAEGE-----HFAVDESLLKQTLVSGAPTGT-----VY 309
QY 168 YNQSDNSTIKLIAMLYTPLRANNAGDDVFTVSCRVLTRPSPDFDFIFLVPPTVESRTKPF 227
DB 310 TSQIPNPLSAVIADSSVPLITDNTATSI-----NPAEPFTFAFNI----- 349
QY 228 TVPILTVEMSNRFPPIPLEKLYTGPSSAFVVPQNGRCTTGGVLLGTTQLSAYNICTFR 287
DB 350 -----TDDVQVRTATL----- 360
QY 288 GDVTHIAGSHDYTNLNASQWNSNYDPTTEIPAPLGPDPFGKIQGMLTQTTREDGSTR-A 346
DB 361 HVTSSAGEAATTNL-----TEDDGSFNWA 385
QY 347 HKATVSTGSHFTPKLGSVQYTTDNTNDPFGQNTKFTPVGVIOGNNHQN-----EPQQ 401
DB 386 LPAADLTGKSWF-----EYTVTATDGF-----NSVTTEPRVTVDGANTDPLRLNLENQ 435
QY 402 WVLPNYSGRT-----GHNVL-----APAV-----APTPEGQL--LFFRSTMP 438
DB 436 WV-----SGTVDVIGASDVFGDKLELLIDDAVATNSSLASAAPTFAEVTQTDVFRNGI- 490
QY 439 GCSGYPNM-----NLDCLLPQEWVQHFCE-----AAPA-----QS 469
DB 491 -LAGGEELRIPOQTYANTETISTEPVLYHINDEGTLTVSVYAGTKAAPEIDLENNDPF 549
QY 470 DVALLRFVNPDTGRVLFECKLHKSQVTVVAHTGPHDLVIPNGYFRFDSWVNPQYTLAPM 529
DB 550 QIRNLRLLLPD-GRTL-----TPAGISDSNAWLN-----M 578
QY 530 GNGAGR 535
DB 579 GDSAGK 584

RESULT 14
US-09-765-272-86
Sequence No. Application US/09765272
Patent No. US20020061545A1
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:

Db 459 REN-----GKIVERTVQGVQFFADSGTDAPSKVLERYVGPE-----FEVPTYYS 505
QY 494 GYVTVAHGTGPHDLVIPP 511
Db 506 NY--QAYDADHPFNPN 521

RESULT 15

US-09-769-787-2
; Sequence 2, Application US/09769787
; Publication No. US20030091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2233
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-2

Query Match 3.9%; Score 111.5; DB 10; Length 2233;
Best Local Similarity 19.9%; Pred. No. 3.3;
Matches 111; Conservative 63; Mismatches 235; Indels 149; Gaps 25;
QY 5 SNDANPSDGSANLPEVNEVNEWALEPVVGAIAAPVAGQQNVDPWIRNPFVQAPGGEF 64
Db 1484 ASDSNPSD-----PVS---NVNDKLISYN--NOPANRW 1511

QY 65 T-----VSPRNPAGEILWSAPLGPDLNPLYLHLMYNGVAGGFVQVILAGNAFTAGKII 120
Db 1512 TNNRTNPASVGVLFQDSGI-----LS--KRSVDNLVGFHED---HGVGVPKSYVI 1559
QY 121 FAAPVPNFFTEGLSPSQVTMPFHIIIVDVRQLEPVL-----IPLPD 160
Db 1560 EYVVGKTVPTAPKNPSFVGNEDHVFNDSANWKFVTNLKAPQKAGEMNHFSFQVETVA 1619

QY 161 VRNPFHYNQSDSTIKLIAMLYTPLRANNAGDDVFTVSCRVLTRPSDFDFIFLVPPTV 220
Db 1620 VRIRMYKADNKRGTSTIEVOIFAKQVAAAKQSGTRIQVDGKDLANFNPDLDIYIL--BSV 1677
QY 221 ESRTKPFPTVILTVEMSNRFPIPLEKLYTGPSSAFVVPONG-----RCTTDGVL 272
Db 1678 DKG-----VPATVA-SVSNGLATVVPVSREGEVPVRIAKAENGDLIGEYRLHFTKDKSL 1731

QY 273 LGTTQLSAYNICTFR--GDV-----THIAGSHDYTMNLASQNSWNSYDPTTEIPAPLG 322
Db 1732 LSHKPVAAVKQARLLQVQALELPTKVPPVFTGKGDETDLTVEM-----EEVPAENL 1785
QY 323 TPDFVGKIQGLMTQITREDCSGTRAHKATVSTGSVHFTPKGSV-----OYTTDTNNDPOT 377
Db 1786 TK-----AQGFTVRGRVLGSNLVABIT-----VRVTDKLGELSDNPYDENSNOQAFAS 1834

QY 378 GQN-----TKFTPVGVITDGNHQNPEQWVLPNYSGRGTGHNVHLAPAVAPTFFGEQLLPF 433
Db 1835 ATNIDKNSHDRVDYLNDDGHSN--RRWT--NWSPTSSNPEVSAGV-----IF 1880
QY 434 RSTMPGCSGYPNMNLDCLLPQEWVQHFCQEAAPQSDVALLRFVNPDTGRVLFECKLHKS 493
Db 1881 REN-----GKIVERTVQGVQFFADSGTDAPSKVLERYVGPE-----FEVPTYYS 1927

QY 494 GYVTVAHGTGPHDLVIPP 511

Db 1928 NY--QAYDADHPFNPN 1943
Search completed: June 1, 2004, 14:04:36
Job time : 37.4562 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 47.0765 Seconds
(without alignments)
3289.030 Million cell updates/sec

Title: US-09-926-799-6
Perfect score: 2896
Sequence: 1 MKMASNDAPNSDGAAGLVP.....VNPFYTLAPMTGNGRRRIQ 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_25Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2896	100.0	548	4 AAB49705	Aab49705 Small rou
2	2813	97.1	548	5 AAU91272	Aau91272 Norwalk v
3	2124	73.3	550	4 AAB49709	Aab49709 Small rou
4	2117.5	73.1	535	4 AAB49707	Aab49707 Small rou
5	2018	69.7	540	4 AAB49706	Aab49706 Small rou
6	1947.5	67.2	539	4 AAB49704	Aab49704 Small rou
7	1931	66.7	542	4 AAB49708	Aab49708 Small rou
8	1678.5	58.0	541	4 AAB49710	Aab49710 Small rou
9	1211	41.8	530	2 AAR50972	Aar50972 Norwalk v
10	1211	41.8	530	7 ADC72176	Adc72176 Norwalk v
11	1209	41.7	530	4 AAB49701	Aab49701 Small rou
12	1178.5	40.7	545	4 AAB49700	Aab49700 Small rou
13	1164	40.2	530	2 AAR57091	Aar57091 Small rou
14	1143.5	39.5	544	4 AAB49703	Aab49703 Small rou
15	1124.5	38.8	546	4 AAB49702	Aab49702 Small rou
16	285.5	9.9	622	4 AAB47045	Aab47045 Feline ca
17	281	9.7	623	4 AAB47044	Aab47044 Feline ca
18	279.5	9.7	668	4 AAB67462	Aab67462 Amino aci
19	277.5	9.6	579	2 AAW08143	Aaw08143 RHDV caps
20	276	9.5	623	4 AAB47043	Aab47043 Feline ca
21	275.5	9.5	547	4 AAM50108	Aam50108 Feline ca
22	275.5	9.5	671	4 AAM50107	Aam50107 Feline ca
23	262	9.0	668	2 AAR10686	Aar10686 Feline ca
24	262	9.0	668	4 AAE04304	Aae04304 Feline ca
25	262	9.0	669	4 AAB67461	Aab67461 Amino aci

ALIGNMENTS

RESULT 1

AAB49705
ID AAB49705 standard; protein; 548 AA.

XX AAB49705;

AC AAB49705;

XX 04-APR-2001 (first entry)

XX Small round structured virus protein SEQ ID 6.

XX Small round structured virus; SRSV; food poisoning.

XX Small round structured virus.

XX WO200079280-A1.

XX 28-DEC-2000.

XX 22-JUN-2000; 2000WO-JP004095.

XX 22-JUN-1999; 99JP-00175928.

XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

XX (DENK-) DENKA SEIKEN KK.

XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

XX WPI: 2001-080848/09.

XX N-PSDB; AAF29146.

XX Kit for the detection and typing of small round-structured virus (SRSV)

XX strains for investigation of food poisoning outbreaks, contains

XX antibodies.

XX Claim 1; Page 52-54; 84pp; Japanese.

XX This invention relates to a kit for the detection and typing of small

XX round structured virus (SRSV) strains. The kit contains antibodies

XX directed against peptides represented in sequences AAB49700 - AAB49710,

XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -

XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is

XX used for detecting and typing strains of SRSV in order to prevent the

XX spread of infection and to examine the epidemiology of outbreaks

XX Sequence 548 AA;

XX SQ

XX Query Match 100.0%; Score 2896; DB 4; Length 548;

XX Best Local Similarity 100.0%; Pred. No. 2.6e-247;

Matches	548;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
y	1	MKQASNDAA	PSNDGAAGLVPEIN	NEFAMALDPVAGAAIAAPLTGQONTIIDPIMN	FVQAP	60			
b	1	MKWASNDAA	PSNDGAAGLVPEIN	NEFAMALDPVAGAAIAAPLTGQONTIIDPIMN	FVQAP	60			
y	61	GGFEFTVSP	RNSPGEVLLNLELGP	INPVLAHARMYNGYAGGFVQVVLACNAFTAGKII	120				
b	61	GGFEFTVSP	RNSPGEVLLNLELGP	INPVLAHARMYNGYAGGFVQVVLACNAFTAGKII	120				
y	121	FAAIPNPFID	NLSAAQITMC	PHVIVDVRLQEPVNLPMDPVRNNEFFHYNQGS	DSRLRLIA	180			
b	121	FAAIPNPFID	NLSAAQITMC	PHVIVDVRLQEPVNLPMDPVRNNEFFHYNQGS	DSRLRLIA	180			
y	181	MLYTPLRANS	GDDVFTVSCRVL	TRPSDFSNFLVPPTVESKTPFTLPILTISEMS	240				
b	181	MLYTPLRANS	GDDVFTVSCRVL	TRPSDFSNFLVPPTVESKTPFTLPILTISEMS	240				
y	241	RFVPVPIES	LHTSPENIVVQC	QNGRVTLDELMTGTTQLLP	SOICAFRGVLTESTGRSDQ	300			
b	241	RFVPVPIES	LHTSPENIVVQC	QNGRVTLDELMTGTTQLLP	SOICAFRGVLTESTGRSDQ	300			
y	301	ADTATPRL	FNYYWHVOLDNL	NGTPYPDAEDI	PGPLGTDPFRGKVFQVASQORNLDSTTTRAH	360			
b	301	ADTATPRL	FNYYWHVOLDNL	NGTPYPDAEDI	PGPLGTDPFRGKVFQVASQORNLDSTTTRAH	360			
y	361	EAKVDTT	AGRTFKPLGSLEI	STSDSDPDQ	NQPTFKPTPVGIGVDNEAEFQWLSLPDYSGOF	420			
b	361	EAKVDTT	AGRTFKPLGSLEI	STSDSDPDQ	NQPTFKPTPVGIGVDNEAEFQWLSLPDYSGOF	420			
y	421	THNNMLA	PAPVAPNPF	GGOLLFP	RSQSPSSGSRNGVLDCLVPQEWQHFVQESAPAQTV	480			
b	421	THNNMLA	PAPVAPNPF	GGOLLFP	RSQSPSSGSRNGVLDCLVPQEWQHFVQESAPAQTV	480			
y	481	ALVRYVNP	DTGKVLFEAKL	HKLGFM	TTIANNGDSPI	TVPPNGYPRFSGSWNPFTTILPMGT	540		
b	481	ALVRYVNP	DTGKVLFEAKL	HKLGFM	TTIANNGDSPI	TVPPNGYPRFSGSWNPFTTILPMGT	540		
y	541	GNRRRIQ	548						
b	541	GNRRRIQ	548						

RESULT 2	
AAU91272	
IID	AAU91272 standard; protein; 548 AA.
XX	
XX	
AC	AAU91272;
XX	
XX	
DT	18-JUN-2002 (first entry)
XX	
DE	Norwalk virus associated polynucleotide #1.
XX	
KW	Nowalk virus; monoclonal antibody; geno group I; geno group II;
KW	immunological detection; food; viral infection.
XX	
OS	Norwalk virus.
XX	
PN	JP2002020399-A.
XX	
PD	23-JAN-2002.
XX	
PF	10-JUL-2000; 2000JP-00208151.
XX	
PR	10-JUL-2000; 2000JP-00208151.
XX	
PA	(OSAP) OSAKA PREFECTURE.
PA	(NISE-) NIPPON SEIBUTSU KAGAKU CENT KK.
PA	(IATR) IATRON LAB INC.
XX	
XX	
DR	WPI; 2002-287412/33.
XX	
PT	A monoclonal antibody useful in the immunological detection and diagnosis

PT	of Norwalk virus infection.
XX	Disclosure; Page 12-13; 24pp; Japanese.
PS	
XX	
CC	The invention describes a monoclonal antibody recognising Norwalk virus, a capsid protein of Norwalk virus, or a common antigen epitope on the capsid protein molecule of geno group I and geno group II. The antibody is useful for immunological detection and quantitative analysis of Norwalk virus in foods and the serum of infected patients. This sequence represents a Norwalk virus associated protein described in the invention
CC	
CC	
XX	Sequence 548 AA;
SQ	
Query Match 97.1%; Score 2813; DB 5; Length 548;	
Best Local Similarity 97.8%; pred. No. 5.8e-240;	
Matches 532; Conservative 6; Mismatches 6; Indels 0; Gaps 0;	
QY	1 MKMASNDAAPSNDGAAGLVPEINNEAMALDPVAGAAATPLTGOONIIDPWIMNFFVQAP 60
Db	1 MKMASRAAPSNDGAAGLVPEINNEAMALDPVAGAAATPLTGOONIIDPWIMNFFVQAP 60
QY	61 GGFEFTVSPRNSGPEVLLNLELGPINPYLAHLARMYNGYAGGFVQVVLGNATAGKII 120
Db	61 GGFEFTVSPRNSGPEVLLNLELGPINPYLAHLARMYNGYAGGFVQVVLGNATAGKII 120
QY	121 FAALPPNFPIDNLSAAQITMCPHVIDVRQLEPYNLPMDPVRRNFFHVNQSGDSRLRLIA 180
Db	121 FAALPPNFPIDNLSAAQITMCPHVIDVRQLEPYNLPMDPVRRNFFHVNQSGDSRLRLIA 180
QY	181 MLYTPLRANNGDDVFTVSCRVLTRPSPDFSNFLVPPTVESKTKPFTLPILTISEMSNS 240
Db	181 MLYTPLRANNGDDVFTVSCRVLTRPSPDFSNFLVPPTVESKTKPFTLPILTISEMSNS 240
QY	241 RFPVPFIESLHSTPTENIVVQCQNGRVTLDDGELMGTTLPSQICAFRGVLFRSTRASDQ 300
Db	241 RFPVPFIESLHSTPTENIVVQCQNGRVTLDDGELMGTTLPSQICAFRGVLFRSTRASDQ 300
QY	301 ADATAPRLFNYYVHVLQDLNLTGTPVDAEDIPGLCTPDFRGKVFVGASQRNLDSTTRAH 360
Db	301 ADIATAPRLFNYYVHVLQDLNLTGTPVDAEDIPGLCTPDFRGKVFVGASQRNPDSTTRAH 360
QY	361 EAKVDVTTAGRTFPKLGSLEISTSDDDFQNGQTKTTPVIGVDNDAEAFQOOSLPDYSGOF 420
Db	361 EAKVDVTTAGRTFPKLGSLEISTSDDFQNGQTRFPGVIGVDNEADFOOSLPDYSGOF 420
QY	421 THNNILAPAVAPNFCEQLLFFRSQLPSSGGHSGNVLDCLVPQEWQHFYQSSAPAQTV 480
Db	421 THNNILAPAVAPNFCEQLLFFRSQLPSSGGHSGNVLDCLVPQEWQHFYQSSAPAQTV 480
QY	481 ALVRYVNPDTGRVLFPAKLHKLGFMTIANKNGDSPITVPNGYFRFPESWNPFTTAPMGT 540
Db	481 ALVRYVNPDTGRVLFPAKLHKLGFMTIANKNGDSPITVPNGYFRFPESWNPFTTAPMGT 540
QY	541 GNGR 544
Db	541 GNGR 544

RESULT 3	
AAB49709	
ID	AAB49709 standard; protein; 550 AA.
XX	
XX	
XX	
XX	AAB49709;
XX	
DT	04-APR-2001 (first entry)
XX	
DE	Small round structured virus protein SEQ ID 10.
XX	
KW	Small round structured virus; SRSV; food poisoning.
XX	
OS	Small round structured virus.
XX	
PN	WO200079280-A1.

XX A monoclonal antibody useful in the immunological detection and diagnosis

```
XX PD 28-DEC-2000.
XX PF 22-JUN-2000; 2000WO-JP004095.
XX PR 22-JUN-1999; 99JP-00175928.
XX PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX PA (DENK-) DENKA SEIKEN KK.
XX PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX DR WPI; 2001-080848/09.
XX DR N-PSDB; AAF29150.
XX XX
XX PT Kit for the detection and typing of small round-structured virus (SRSV)
XX PT strains for investigation of food poisoning outbreaks, contains
XX PT antibodies.
XX XX
XX PS Claim 1; Page 62-64; 84pp; Japanese.
XX CC This invention relates to a kit for the detection and typing of small
XX CC round structured virus (SRSV) strains. The kit contains antibodies
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX CC used for detecting and typing strains of SRSV in order to prevent the
XX CC spread of infection and to examine the epidemiology of outbreaks
XX XX
XX SQ Sequence 550 AA;
XX
Query Match 73.3%; Score 2124; DB 4; Length 550;
Best Local Similarity 71.4%; Pred. No. 6e-179;
Matches 396; Conservative 72; Mismatches 75; Indels 12; Gaps 4;
QY 1 MKMASNDAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQONIIDPWIMNFVQAP 60
DB 1 MKMASNDAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQONIIDPWIMNFVQAP 60
QY 61 GGEFTVSPRNSGEGVLLNLELGPENPYLAHLARMYNGVAGFVQVVLGNAFTAGKII 120
DB 61 GGEFTVSPRNSGEGVLLNLELGPENPYLAHLARMYNGVAGFVQVVLGNAFTAGKII 120
QY 121 FFAAIPNPFIDNLSAAQITMCPHVIVDVRLQLEPVLNLPMDVRNPFHYNQGSDSLRLIA 180
DB 121 FFAAIPNPFIDNLSAAQITMCPHVIVDVRLQLEPVLNLPMDVRNPFHYNQGSDSLRLIA 180
QY 181 MLYTPLRANSGDDVFTVSCRVLTSPDFSNFLVPTVESKTKPFTLPILTISEMNS 240
DB 181 MLYTPLRANSGDDVFTVSCRVLTSPDFSNFLVPTVESKTKPFTLPILTISEMNS 240
QY 241 RFPVPTESLHTSPENIVVQCONGRVTLTGELMGTTQLLPQICAFRGLVTRSTGRASDQ 300
DB 241 RFPVPTESLHTSPENIVVQCONGRVTLTGELMGTTQLLPQICAFRGLVTRSTGRASDQ 300
QY 301 ADTATPRLNYYWHVQDLNLTGTPYDPAEDIPGLTDPDFRGKVFQVQASQRLNLS - 355
DB 301 ADTATPRLNYYWHVQDLNLTGTPYDPAEDIPGLTDPDFRGKVFQVQASQRLNLS - 355
QY 300 TD-SPOARNHPLHVQVKLDTQYDPTDDIPAVLGAIDFKGVFGVQASQRDVSGQEQG 358
DB 300 TD-SPOARNHPLHVQVKLDTQYDPTDDIPAVLGAIDFKGVFGVQASQRDVSGQEQG 358
QY 356 --TTTRAHEAKVDTAGRTFKLGSLEISTDDDFONQTKTTPVIGVGNDAEFQOWSL 413
DB 356 --TTTRAHEAKVDTAGRTFKLGSLEISTDDDFONQTKTTPVIGVGNDAEFQOWSL 413
QY 359 HYATRAHEAHIDTTPKYAPKLGTLIKSGSDDFNTNQIRFTPVGMGDN--WQWEL 415
DB 359 HYATRAHEAHIDTTPKYAPKLGTLIKSGSDDFNTNQIRFTPVGMGDN--WQWEL 415
QY 414 PDYSGQFTHMNLAPAVAFNPFGEOLLRPSOLPSSGGRSNGVLCLVPQEWQHFYQES 473
DB 414 PDYSGQFTHMNLAPAVAFNPFGEOLLRPSOLPSSGGRSNGVLCLVPQEWQHFYQES 473
QY 416 PDYSGRLTUNMLAPAVSFPPGKRLIFRSPVAGGYSYIIDLIPQEWQHFYQEA 475
DB 416 PDYSGRLTUNMLAPAVSFPPGKRLIFRSPVAGGYSYIIDLIPQEWQHFYQEA 475
QY 474 APAQTOVALRVYVNPDTGKVLFEAKLHKLGMFTIANNGDSPTTPPNPGYFRPESWNPY 533
DB 474 APAQTOVALRVYVNPDTGKVLFEAKLHKLGMFTIANNGDSPTTPPNPGYFRPESWNPY 533
QY 476 AFSQSAVALRVYVNPDTGKVLFEAKLHKLHREGFLTVANGCNPVIVPNPGVFRFANGNQFY 535
DB 476 AFSQSAVALRVYVNPDTGKVLFEAKLHKLHREGFLTVANGCNPVIVPNPGVFRFANGNQFY 535
QY 534 TLAPMGTCNGRRRIQ 548
DB 534 TLAPMGTCNGRRRIQ 548
```

```
DB 536 TLAPMGTCNGRRRAQ 550
DB
RESULT 4
AAB49707
ID AAB49707 standard; protein; 535 AA.
XX
XX AAB49707;
AC
XX
XX 04-APR-2001 (first entry)
XX
XX Small round structured virus protein SEQ ID 8.
XX
XX Small round structured virus; SRSV; food poisoning.
XX
XX Small round structured virus.
XX
XX WO200079280-A1.
XX
XX 28-DEC-2000.
XX
XX 22-JUN-2000; 2000WO-JP004095.
XX
XX 22-JUN-1999; 99JP-00175928.
XX
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX PA (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI; 2001-080848/09.
XX DR N-PSDB; AAF29148.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX PT strains for investigation of food poisoning outbreaks, contains
XX PT antibodies.
XX
XX Claim 1; Page 57-59; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
XX CC round structured virus (SRSV) strains. The kit contains antibodies
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX CC used for detecting and typing strains of SRSV in order to prevent the
XX CC spread of infection and to examine the epidemiology of outbreaks
XX XX
XX SQ Sequence 535 AA;
XX
Query Match 73.1%; Score 2117.5; DB 4; Length 535;
Best Local Similarity 70.5%; Pred. No. 2.2e-178;
Matches 387; Conservative 74; Mismatches 73; Indels 15; Gaps 4;
QY 1 MKMASNDAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQONIIDPWIMNFVQAP 60
DB 1 MKMASNDAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQONIIDPWIMNFVQAP 60
QY 61 GGEFTVSPRNSGEGVLLNLELGPENPYLAHLARMYNGVAGFVQVVLGNAFTAGKII 120
DB 61 GGEFTVSPRNSGEGVLLNLELGPENPYLAHLARMYNGVAGFVQVVLGNAFTAGKII 120
QY 61 NGEFTVSPRNSGEGVLLNLELGPENPYLAHLARMYNGVAGFVQVVLGNAFTAGKIV 120
DB 61 NGEFTVSPRNSGEGVLLNLELGPENPYLAHLARMYNGVAGFVQVVLGNAFTAGKIV 120
QY 121 FFAAIPNPFIDNLSAAQITMCPHVIVDVRLQLEPVLNLPMDVRNPFHYNQGSDSLRLIA 180
DB 121 FFAAIPNPFIDNLSAAQITMCPHVIVDVRLQLEPVLNLPMDVRNPFHYNQGSDSLRLIA 180
QY 181 MLYTPLRANSGDDVFTVSCRVLTSPDFSNFLVPTVESKTKPFTLPILTISEMNS 240
DB 181 MLYTPLRANSGDDVFTVSCRVLTSPDFSNFLVPTVESKTKPFTLPILTISEMNS 240
QY 241 RFPVPTESLHTSPENIVVQCONGRVTLTGELMGTTQLLPQICAFRGLVTRSTGRASDQ 300
DB 241 RFPVPTESLHTSPENIVVQCONGRVTLTGELMGTTQLLPQICAFRGLVTRSTGRASDQ 300
QY 241 RFPVPTESLHTSPENIVVQCONGRVTLTGELMGTTQLLPQICAFRGLVTRSTGRASDQ 300
DB 241 RFPVPTESLHTSPENIVVQCONGRVTLTGELMGTTQLLPQICAFRGLVTRSTGRASDQ 300
```



```
PS Claim 1; Page 50-52; 84pp; Japanese.
XX This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX Sequence 539 AA;

Query Match 67.2%; Score 1947.5; DB 4; Length 539;
Best Local Similarity 66.2%; Pred. No. 2.5e-163;
Matches 365; Conservative 73; Mismatches 94; Indels 19; Gaps 4;

QY 1 MKMASNDAAPSNDGAGLVPEINNEAMALDPVAGAAIAAPLTGQOQNIIDPWIMNMFVQAP 60
DB 1 MKMASNDANPSGDTANLVPEVNEVMALEPVVGAIAAPVAGQOQNVDPWRNMFVQAP 60

QY 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGFEVQVVLGNAGTAAGKII 120
DB 61 GGEFTVSPRNPAGEILWSAPLGPDLNPLYSLHARMYNGYAGFEVQVVLGNAGTAAGKII 120

QY 121 FFAAIPNPIDNLSAAQITMCPHVIVDVRQLEPVMPLPMDVNRNFFHYNQGSDSLRLIA 180
DB 121 FFAAIPNPIDNLSAAQITMCPHVIVDVRQLEPVMPLPMDVNRNFFHYNQGSDSLRLIA 180

QY 181 MLYTPLRANNSGDDVFTVSCRVLTRPSPDFSNFLVPPPTVESKTKPFTLPILTISEMSNS 240
DB 181 MLYTPLRANNSGDDVFTVSCRVLTRPSPDFSNFLVPPPTVESKTKPFTLPILTISEMSNS 240

QY 241 RFPVPIESLHTSPTEINIVVQCONGRVTLDELGMGTTLQLPSQICAFRGVLTSTRSDQ 300
DB 241 RFPVPIESLHTSPTEINIVVQCONGRVTLDELGMGTTLQLPSQICAFRGVLTSTRSDQ 300

QY 301 ADTATPRLFNYVHVOLNLTGTPDPAEDIPGLGTDPFRGKVFGVASQRLN-DSTTRA 359
DB 301 ADTATPRLFNYVHVOLNLTGTPDPAEDIPGLGTDPFRGKVFGVASQRLN-DSTTRA 359

QY 360 HEAKVDTTAGRTPKLGSLEISTD-SDPDFQOQTKFTPVGI---GVNEAEFQOQSLPD 415
DB 360 HEAKVDTTAGRTPKLGSLEISTD-SDPDFQOQTKFTPVGI---GVNEAEFQOQSLPD 415

QY 347 HKATVSTGSHVHTPKLGSVQYITDNNDFQTQNTKFTPVGIQDGNHNPQOQVLPN 406
DB 347 HKATVSTGSHVHTPKLGSVQYITDNNDFQTQNTKFTPVGIQDGNHNPQOQVLPN 406

QY 416 YSQQFTHNNMLAPAVAPNPPGQQLFFRSQPSGSRNSGVLDCLVPQEWVQHFYQESAP 475
DB 407 YSGRTGNVHLAPAVAPNPPGQQLFFRSQPSGSRNSGVLDCLVPQEWVQHFYQESAP 466

QY 476 AQOVALRVYNPDGKVLFEAKLHKLGMFTIANGSDSPITVPPNGYFRFESWNPFFYL 535
DB 467 AQSDVALLRfVNPDTGRVLFECKLHKSGYVTVTAHTGPHDLVIPPNGYFRFDSWVNFYTL 526

QY 536 APWGTGNGRRR 546
DB 527 APMGNGAGRRR 537

RESULT 7
AAB49708
ID AAB49708 standard; protein; 542 AA.
XX
XX AAB49708;
AC
DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 9.
XX
XX Small round structured virus; SRSV; food poisoning.
XX
XX Small round structured virus.
XX
XX Small round structured virus.
XX
XX WO200079280-A1.
DB
```

```
PD 28-DEC-2000.
XX
XX 22-JUN-2000; 2000WO-JP004095.
XX
XX 22-JUN-1999; 99JP-00175928.
XX
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI: 2001-080848/09.
XX N-PSDB; AAF29149.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX strains for investigation of food poisoning outbreaks, contains
XX antibodies.
XX
XX Claim 1; Page 59-61; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks
XX Sequence 542 AA;

Query Match 66.7%; Score 1931; DB 4; Length 542;
Best Local Similarity 65.7%; Pred. No. 7.3e-162;
Matches 364; Conservative 70; Mismatches 102; Indels 18; Gaps 5;

QY 1 MKMASNDAAPSNDGAGLVPEINNEAMALDPVAGAAIAAPLTGQOQNIIDPWIMNMFVQAP 60
DB 1 MKMASNDAAPSNDGAGLVPEINNEAMALDPVAGAAIAAPLTGQOQNIIDPWIMNMFVQAP 60

QY 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGFEVQVVLGNAGTAAGKII 120
DB 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGFEVQVVLGNAGTAAGKII 120

QY 121 FFAAIPNPIDNLSAAQITMCPHVIVDVRQLEPVMPLPMDVNRNFFHYNQGSDSLRLIA 180
DB 121 FFAAIPNPIDNLSAAQITMCPHVIVDVRQLEPVMPLPMDVNRNFFHYNQGSDSLRLIA 180

QY 181 MLYTPLRANNSGDDVFTVSCRVLTRPSPDFSNFLVPPPTVESKTKPFTLPILTISEMSNS 240
DB 181 MLYTPLRANNSGDDVFTVSCRVLTRPSPDFSNFLVPPPTVESKTKPFTLPILTISEMSNS 240

QY 241 RFPVPIESLHTSPTEINIVVQCONGRVTLDELGMGTTLQLPSQICAFRGVLTSTRSDQ 300
DB 241 RFPVPIESLHTSPTEINIVVQCONGRVTLDELGMGTTLQLPSQICAFRGVLTSTRSDQ 300

QY 301 ADTATPRLFNYVHVOLNLTGTPDPAEDIPGLGTDPFRGKVFGVASQRLN-DSTTRA 359
DB 301 ADTATPRLFNYVHVOLNLTGTPDPAEDIPGLGTDPFRGKVFGVASQRLN-DSTTRA 359

QY 360 HEAKVDTTAGRTPKLGSLEISTD-SDPDFQOQTKFTPVGI---GVNEAEFQOQSLPD 416
DB 360 HEAKVDTTAGRTPKLGSLEISTD-SDPDFQOQTKFTPVGI---GVNEAEFQOQSLPD 416

QY 349 KNAKIGIYSTSGRTPKIGSITGLHSITEHVHPNQOQSRFTPVGVAVDENTPPQOQVLPY 408
DB 349 KNAKIGIYSTSGRTPKIGSITGLHSITEHVHPNQOQSRFTPVGVAVDENTPPQOQVLPY 408

QY 417 SQQFTHNNMLAPAVAPNPPGQQLFFRSQPSGSG--RSNGVLDCLVPQEWVQHFYQESA 474
DB 409 AGSLALNTLAPAVAPNPPGQQLFFRSQPSGSGSG--RSNGVLDCLVPQEWVQHFYQESA 468

QY 475 PAQOVALRVYNPDGKVLFEAKLHKLGMFTIANGSDSPITVPPNGYFRFESWNPFFYL 534
DB 469 PSQADVALRVYNPDGKVLFEAKLHKLGMFTIANGSDSPITVPPNGYFRFESWNPFFYL 528

QY 535 LAPMGGTGNGRRRIQ 548
DB 529 LAPMGGTGNGRRRIQ 542
```

```

Db 300 -----YRDMETNTDGPIDTPTPGIPSGDFQILFGVASQRNKEQMPA 348
QY 357 TRAHEAKVDTTAGRTFKLGLSLEISDSDDFDQ-NQTKFTFVG-----IGVDNEAEFQ 409
Db 349 TRAHEAIINTGGDHLCPQISSEIYLTSPNLRCTNPNQLPQSGLRGTILIRSDN-----403
QY 410 QMSLPDYSQGFTHNNLAPAVAPNPPGQQLLFFRSQPSGSGRSNGVLDCILVPGQEWQHF 469
Db 404 -----GHCHDMVGTSTPTTPWQWRRCSSRGSSGHRYVPVVMNRVTMIVLS 454
QY 470 YOBAPAPQTV-----ALVRYVNPDTGKVLFAKLHKLGFMTIANNDSPIVPPNG 521
Db 455 HKSGFSTIRKLPQLNLRWPLIRFINPOTGRVLFEARLHKQGFITVAHTGDNPIVMPNG 514
QY 522 YRFESWVNPFFYTLAPMGTGNGRRRIQ 548
Db 515 YRFEAWNQFYSLAPVGTGKRRRVQ 541

RESULT 9
AAR50972
ID AAR50972 standard; protein; 530 AA.
XX
AC AAR50972;
XX
DT 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 05-OCT-1994 (first entry)
XX
DE Norwalk virus strain 8FIIa protein (encoded by ORF2).
XX
KW Norwalk virus; pathogen; acute gastroenteritis; food poisoning;
KW seafood contamination; diagnostic assay; calicivirus; small round virus.
XX
OS Norwalk virus; (strain 8FIIa).
XX
PN WO9405700-A2.
XX
PD 17-MAR-1994.
XX
PF 07-SEP-1993; 93WO-US008447.
XX
PR 07-SEP-1992; 92US-00941365.
XX
PA (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI Matson DO, Estes MK, Jiang X, Graham DY;
XX
WPI; 1994-101125/12.
DR N-PSDB; AAQ56826.
XX
PT DNA from Norwalk and related viruses - used for preparing prods. for use
PT in diagnostic assays, detection and vaccines for Norwalk and related
PT viruses.
XX
PS Claim 14; Page 68-70; 156pp; English.
XX
The Norwalk virus was isolated from stool samples from adult volunteers
infected with safety tested Norwalk virus strain 8FIIa. The coding
sequence is useful for the design of probes for use in diagnostic assays
for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct
FN field.) (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 530 AA;
Query Match 41.8%; Score 1211; DB 2; Length 530;
Best Local Similarity 45.4%; Pred. No. 4.1e-98;
Matches 260; Conservative 74; Mismatches 161; Indels 78; Gaps 11;

QY 1 MKVASNDAAAPSDGAG---LVPEIN-NEAMALDPVAGAAIAAPITGQONIIDPWIMNPF 56
Db 1 MMASKDQATSSVDGASGAGQVLEFVNASDPLANDPVAGSLAVATAGVNPIDPWIMNPF 60

This invention relates to a kit for the detection and typing of small
round structured virus (SRSV) strains. The kit contains antibodies
directed against peptides represented in sequences AAB49700 - AAB49710,
which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
AAF20151 represent cDNA encoding the strain specific proteins. The kit is
used for detecting and typing strains of SRSV in order to prevent the
spread of infection and to examine the epidemiology of outbreaks
XX
SQ Sequence 541 AA;
Query Match 58.0%; Score 1678.5; DB 4; Length 541;
Best Local Similarity 58.9%; Pred. No. 1.7e-139;
Matches 334; Conservative 67; Mismatches 121; Indels 45; Gaps 7;

QY 1 MKVASNDAAAPSDGAGLVPEINNEAMALDPVAGAAIAAPITGQONIIDPWIMNPFVQAP 60
Db 1 MKVASNDAAAPSDGAGLVPEINNEAMALDPVAGAAIAAPITGQONIIDPWIMNPFVQAP 60
QY 61 GGEFTVSPNSPCEVLLNLELPEINPYLAHLMVNGVAGGFEVQVVLGNAFTAGKII 120
Db 61 AGEFTVSPNSPGEIILLDELGPDLNPNYLAHLMVNGHAGGVEQIVLGNAGFTAGKII 120
QY 121 FRAIPNFPIIDNLAAQITMCWHIVDVVRQLEFVNLPMFDPVRNPNFHYNQSGDSRLRLIA 180
Db 121 FRAIPFGFVYENLSPQITMCWHIVDVVRQLEFVNLPMFDPVRNPNFHYNQSGDRLRLVA 180
QY 181 MLYTPLRANNSGDDVTVSCRVLTRSPDFSNFLVPTVESKTKPFTLPILITISEMSN 240
Db 181 MLYTPLRANNSGDDVTVSCRVLTRSPDFSNFLVPTVESKTKPFTLPILITISEMSN 240
QY 241 RFVPVIESLTSPTENIVWQCONGRVTLDGELMGTQLLPSQICAFRG-VLTRSTSRASD 299
Db 241 RFVPVDMVTARNENQVQPNQGRVTLDGELMGTQLLPSQICAFRG-VLTRSTSRASD 299
QY 300 QADTATPRLFNYWHVQDLNGLNGTPVDPDAEDIFGPIGTDFRGKVFVGASQRNLDLS---T 356

```

[illegible]

PT Norwalk virus.

XX

XX Example 4; SEQ ID NO 3; 45pp; English.

XX

This invention relates to a novel method of inducing an immune response in an individual against Norwalk virus and non-Norwalk virus agents, by orally or parenterally administering an immunogen recombinantly expressed or synthesised from a cDNA of Norwalk virus given in the specification. Norwalk virus is one of the most important viral pathogens, causing acute gastroenteritis. The invention may be used for the development of compounds with virucidal activity or an antiviral vaccine. The present sequence is the amino acid sequence of a protein encoded by the Norwalk virus genome of the invention.

XX

XX Sequence 530 AA;

XX

Query Match 41.8%; Score 1211; DB 7; Length 530;
Best Local Similarity 45.4%; Pred. No. 4.1e-98;
Matches 260; Conservative 74; Mismatches 161; Indels 78; Gaps 11.

QY 1 MKMASNDAAPSNDGAAG---LVPEIN-NEAMALDPVAGAAIAAPLTGQQNIIDPWNNF 56
DB 1 MMWASKDATSSVDGASGAGQLVPEVNASDPLAMPVAGSSTAVATAGQVNPIDPWNNF 60

QY 57 VQAPGGEFTVSPNSGCEVLLNLELGPENIPYLALHARMYNGVAGGFEVQVVLGAPTA 176
DB 61 VQAPOGEFTISPNTGDLVFDLSLGPHLNPFLLHLSQMYNGVGNMVRIMLAGNAFTA 120

QY 117 GKIIIPAAIPNPIDNLNSAAQITMCPHVIVDVRQLEPNVLPMPDVRRNPFHYNQGSRL 176
DB 121 GKIFVSCIPPGFESHNLITAAQLFPHVITADVRLTLOPIEVLDERVNLPHNDRNQTM 180

QY 177 RLIAMLYTPLRANNSGDDVFTVSCRVLTSPDFSENFVLPPTVESKTPFTLPILTISE 236
DB 181 RLVCMLYTPLRCTGGTGSFVWAGRVMTCPSPDFNEDLVLPPTVEQKTRPFTLPNPLSS 240

QY 237 MNSRPPVPIESLHTSPNTENIVVQCONGRVITDGEMLGTQLLPQICAPRGVLTSTSR 296
DB 241 LNSRAPLPTISSMGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRG-----TSN 295

QY 297 ASDQADATPRLFNYYHVLQNLNACTPYDPAEDIPGLGTDPFRG-----KVFGVAS 349
DB 296 GT-----VINLTEDGTTPHPFEG-PAPIGFDPDGGCDWHINNTQFGHSS 339

QY 350 QRNLDDSTTRAHEAKVDTTAGRTTPKLGSLSEISTDSDDFDQNPQTKFTPVGIVDN----- 404
DB 340 QTYQD-----VDTTPTDTPVPLGSIQAN-----GIGSGNYGVGL 373

QY 405 -----EAEFOQMSLPDYSGQPTTHNNILAPAVAPNFPGEQLLFFRSQLPSSGGRSN 454
DB 374 SMISPPSPHSQSDVLWKIPNYGSSITTEATHLAPSVPFGFGEVLVFPNMSKMPGPGAYN- 432

QY 455 GVLDCLVPOEWQHFQESAPAQTVALRYVNPDTGKVLFEAKLHKLGFMTIANNQDS- 513
DB 433 --LPCLLPQEYISHLASEQAPTVEAALHYVDPDTGRNLGFBFKAYPDGFLTCVPGNASS 490

QY 514 -PTTPVPPNGVYRFESVMNPFYTLAPMGTCNGRR-545
DB 491 GPOQLFINGVFVFSWVSRYQLKPVGTASSAR 523

RESULT 11
AAB49701
ID AAB49701 standard; protein; 530 AA.
XX
XX AAB49701;
XX
XX 04-APR-2001 (first entry)
XX
XX Small round structured virus protein SEQ ID 2.
XX
XX Small round structured virus; SRSV; food poisoning.
XX


```

Db 240 NTLNSRVPSLKMVMVRDQHQVQFQNGRVTLGQQLGTTPTASQCLCKIRGSVFHAN 299
Qy 295 SRASQADTATPRLFNYYWHVOLDNLNGTPYDPAEDIPGLGTPDFRGKV-----FGVAS 349
Db 300 GG-----NGY---NLTELDGSPYHAFES-PAPIGFPDL-GECDWHMEASPTT 341
Qy 350 QRNLSTTTRAHAKVDTTAGRTPKLGSLEISTDS-DFDQNPYKF---TPVGIGVDNE 405
Db 342 QNTGDIKQINVKQESA---FAPHLGTIQADGLSDVSVNTMIKLGWSPVSDG--HR 396
Qy 406 AEFQWMLPDYSGQFTHNMNLAPAVAPNPPGQOLLFFRSQSSGSRNGVLDCILVROEW 465
Db 397 GVDVFWVIRYSGTITEAQAAPPYPPGFGGAIVFFMSDFPIAHTGNGLSVPCITPOEF 456
Qy 466 VOHFYQESAPQTOVALVRYNPNPTGKLVFEAKHLKLGFMFTIANN--GDSPTITPPNGYF 523
Db 457 VTHFVNEQAPTRGEAALLHYLDPDTHRLNGEFLYPEGFMTCVPSNGSGTGPTLPINGVF 516
Qy 524 RPSWNPFFYTLAPMGNGRRRI 547
Db 517 VFVSWVSRYQLKPVGTAGPACRL 540

```

RESULT 13

AA57091
ID AARS7091 standard; protein; 530 AA.

AA57091;
AC AARS7091;

27-AUG-2003 (revised)

25-MAR-2003 (revised)

05-OCT-1994 (first entry)

Small round virus SRSV/KY/89 capsid protein.

pathogen; acute gastroenteritis; food poisoning; seafood contamination;

diagnostic assay; human calicivirus; small round virus; SRSV; KY89;

Norwalk virus; capsid protein.

Small round structured virus.

WO9405700-A2.

17-MAR-1994.

07-SEP-1993; 93WO-US008447.

07-SEP-1992; 92US-00941365.

(BAYU) BAYLOR COLLEGE MEDICINE.

Matson DO, Estes MK, Jiang X, Graham DY;

WPI; 1994-101125/12.

N-PSDB; AAQ56832.

DNA from Norwalk and related viruses - used for preparing prods. for use
in diagnostic assays, detection and vaccines for Norwalk and related
viruses.

Example 7; Fig 13a; 156pp; English.

The known sequence for Norwalk virus was used to obtain the sequence of
other Norwalk-related viruses such as SRSV/KY/89, an agent from a stool
from an outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide
cDNA sequence includes part of the polymerase region and the capsid
region of the genome; the deduced amino acid sequences are AARS7092 and
AARS7091, respectively. Expression of fragments and derivs. of Norwalk-
related viruses permits development of diagnostic assays to detect
antibodies, antigens, viral genetic material or antivirals. (Updated on
25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS
field.)

```

XX SQ Sequence 530 AA;
Query Match 40.2%; Score 1164; DB 2; Length 530;
Best Local Similarity 44.4%; Pred. No. 5, 9e-94;
Matches 255; Conservative 74; Mismatches 165; Indels 80; Gaps 13;
Qy 1 MMSASNDAAPSNDGAG---LVPEIN-NEAMALDPVAGAAIAAPLTGQNIIDPIMMNP 56
Db 1 MMSASKDATSSVDGASASVQLVPEVNASDPLAMDPAVAGSTAVATAGQVNPIDPWIINF 60
Qy 57 VOAPGGEFTVSPRNSPGEVLNLELGPINPVLALARMYNGYAGGFFVOVVLGNAFTA 116
Db 61 VOAPGGEFTISPNNTPGDVLFDLSLGPHLNPFLLHLSQYNGWGMVRIMLAGNAFTA 120
Qy 117 GKIPAAIPPPFPIDNLSAAQITMCCHPVIIVDVRQLEPNLMPDPVRRNFFHYNOGSDSRL 176
Db 121 GKIIVSCIPPGSGQSLTIAQATLFFHVIADVRLDPIEVLDPVRLVFNHNDNRNQOTM 180
Qy 177 RLIALMYTPLRANNSGDDVFTVSCRVLTRPSDFSNFLVPPTVESKTKPFTLPILTISE 236
Db 181 RLVCMLYTPLSTGGTGDGDFVAVGRVMTCPSPDFNFLVPPTVBQKTRPFTLPNPLSS 240
Qy 237 MNSRPPVPVPIESLHTSPTEINIVQCGRVTLDGELMGTTLQLLPSQICAPRGVLTSTRSR 296
Db 241 LNSRAPLPISGMGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRG-----TSN 295
Qy 297 ASDQADTATPRLFNYYWHVQLDNLNGTPYDPAEDIPGLGTPDFRG-----KVFGVAS 349
Db 296 GT-----VINLTEDGTFFHPEG-PAPIGFPDLGGCDWHNMTOFGHSS 339
Qy 350 QRNLSTTTRAHAKVDTTAGRTPKLGSLEISTDSDFDQNPYKFPTVIGVDN-----404
Db 340 QTQVD-----VDTTPTDSVPHLGSIQAN-----GIGSGNYIGVL 373
Qy 405 -----EAEFOQWSLPDYSQOFTNMNLAPAV-APNPFGEQLLFFRSQSPSSGGRS 453
Db 374 SWSPPSPHSPSGQVDMKIPNYGSSITEATHLAPSVSPGF-GEYLVFFMSKIPGPGDS 432
Qy 454 NGVLDCILVPOQWVQHFYQESAPQTOVALVRYNPNPTGKLVFEAKHLKLGFMFTIANN 513
Db 433 ---LPCLLPQGYIISHLASEQAPTVEGELLHYVDDTDNRNLGEFKAYPDGFLTCVPNGAS 489
Qy 514 --PITVPNGYFRFESWYNPFYTLAPMGNGRR 545
Db 490 SGPQQLPINGVFVFWVSWSRYQLKPVGTASTAR 523
RESULT 14
AAB49703
ID AAB49703 standard; protein; 544 AA.
XX AC AAB49703;
XX DT 04-APR-2001 (first entry)
XX DE Small round structured virus protein SEQ ID 4.
XX KW Small round structured virus; SRSV; food poisoning.
XX OS Small round structured virus.
XX PN WO200079280-A1.
XX PD 28-DEC-2000.
XX PF 22-JUN-2000; 2000WO-JP004095.
XX PR 22-JUN-1999; 95JP-00175928.
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
XX

```

Wed Jun 2 09:13:30 2004

PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 XX WPI; 2001-080848/09.
 DR N-PSDB; AAF29144.
 XX
 DR
 XX
 DR
 XX
 PT Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX
 PS Claim 1; Page 47-49; 84pp; Japanese.
 XX
 XX This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX
 SQ Sequence 544 AA;
 Query Match 39.5%; Score 1143.5; DB 4; Length 544;
 Best Local Similarity 44.5%; Pred. No. 4e-92;
 Matches 252; Conservative 86; Mismatches 183; Indels 45; Gaps 15;
 QY 1 MKVASNDAAPSNDGAAG---LVPEINN-EMALDPVAGAAIAAPLTGQONIIDPWIMNMF 56
 Db 1 MMASKDATSADGATGAGQVPEVNTADPIIDPVAGSSTALATAGQVNLIDPWIMNF 60
 QY 57 VQAPGGFTVSPRNSPGEVLNLELGPENPYLAHLARMYNGYAGGFEVQVVLGNAFTA 116
 Db 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNFPLSHLSQMYNGWGNRVRLVLAGNAFTA 120
 QY 117 GKIIFAAIPNFPIDNLSAAQITMCPHVIDVROLEPVNLMPDPVNRNFFHYNQGSRL 176
 Db 121 GKVIICCPVPGFQRTSLIAQATLFPHVIAVRLDPEVPLEDRVRLVYH-NNDTQPTM 179
 QY 177 RLIAMLYTPLRAN--NSGDDVFTVSCRVLTSPDPSFNFLVPTVESKTKPFTLPILTI 234
 Db 180 RLLCMLYTPLTGASGTDTSFVAGRVLTCPGDFNLFVLPVPTVQKTRFTVFNIP 239
 QY 235 SEMSNRPVPIESLHTSPENIVVQCGRVTLGDLGMLGTTQLLPSQICAFRGVLTST 294
 Db 240 KYLSNSRIPNPIEGNSLSPDOTQNVQNGRCTIDGQPLGTPVSVLSQCKFRGRI---- 295
 QY 295 SRASDAQDTATPLRPNFYHVVOLDNLNGTPYDPAEDIPGLTDPDFRGKVFVASQRNLD 354
 Db 296 -----TSQORVLN-----LTELDSGF-MAFAAPAPAGFPDLGSCDWHIEMSKIPN 340
 QY 355 STTRAHEA---KYDTTAGRTPKLSL---EISTSDDFDQNOPTKFTPVGIGVDNEAE 407
 Db 341 SSTQNNPIVTSKVNKSQOQVPHLSLITLDENSVSGDYIGTIQWTS-PPSDSG--GAN 396
 QY 408 FQOWSLPDYSGQFTHNMMLAPAVNFFEQQLLFFRSQLP---SSGRSGNVLDCLVPOE 464
 Db 397 TNFWKIPDYGSSIAEASQAPAVYPGGENIVYFEMASIPGNOSG--SPNLVPCLLPOE 454
 QY 465 WYOHFQESAPAQTVQALVRYNPDTKVLFEAKHLKLGFWMTIANGDS--PITVPPNGY 522
 Db 455 YITHFTISEQAPIQGEAALHYDPTNRNLGFEKLYPGGYLTCVPSNSSSTGPOQLPDGV 514
 QY 523 PFESWNPFTYLA PMGT--GNGRRRI 547
 Db 515 FVFASWWSRFYQLKPVGTAGPARGRL 540
 RESULT 15
 AAB49702
 ID AAB49702 standard; protein; 546 AA.
 XX
 AC AAB49702;
 XX
 DT 04-APR-2001 (first entry)

XX Small round structured virus protein SEQ ID 3.
 DE
 XX Small round structured virus; SRSV; food poisoning.
 KW
 XX Small round structured virus.
 OS
 XX
 XX WO200079280-A1.
 PN
 XX
 PD 28-DEC-2000.
 XX
 XX 22-JUN-2000; 2000WO-JP004095.
 XX
 PF 22-JUN-1999; 99JP-00175928.
 PR
 XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.
 PA
 XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 PI
 XX WPI; 2001-080848/09.
 XX N-PSDB; AAF29143.
 DR
 XX Kit for the detection and typing of small round-structured virus (SRSV)
 XX strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 PT
 XX Claim 1; Page 45-47; 84pp; Japanese.
 PS
 XX This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 CC
 XX Sequence 546 AA;
 SQ
 Query Match 38.8%; Score 1124.5; DB 4; Length 546;
 Best Local Similarity 43.0%; Pred. No. 1.9e-90;
 Matches 250; Conservative 80; Mismatches 182; Indels 69; Gaps 17;
 QY 1 MKMASNDAAPSNDGAAG---LVPEINN-EMALDPVAGAAIAAPLTGQONIIDPWIMNMF 56
 Db 1 MMASKDAPQADGASGAGQVPEVNTADPLPMEPVAGPTTAVATAGVNMIDPWIMNMF 60
 QY 57 VQAPGGEFTVSPRNSPGEVLNLELGPENPYLAHLARMYNGYAGGFEVQVVLGNAFTA 116
 Db 61 VQSPQGEFTISPNNTPGDILFDLQGLPHLNFPLSHLSQMYNGWGNRVRLVLAGNAFSA 120
 QY 117 GKIIFAAIPNFPIDNLSAAQITMCPHVIDVROLEPVNLMPDPVNRNFFHYNQGSRL 176
 Db 121 GKIIIVCCVPPGFTSSSLTIAQATLFPHVIAVRLDPEVPLEDRVRLVYHNTDQPT-M 179
 QY 177 RLIAMLYTPLR--ANNSGDDVFTVSCRVLTSPDPSFNFLVPTVESKTKPFTLPILTI 234
 Db 180 RLVCMLYTPLTGCGSGNSDSFVAGRVLTAPSDSFSLFVLPPTIEQKTRFTVPTPL 239
 QY 235 SEMSNRPVPIESLHTSPENIVVQCGRVTLGDLGMLGTTQLLPSQICAFRGVLTST 294
 Db 240 QTLNSRFPFLSQMILSPDASQVQFQNGRCLIDGLLGTTPATSGQLFVRVGKINQGA 299
 QY 295 SRASDAQDTATPLRPNFYHVVOLDNLNGTPYDPAEDIPGLTDPDFRGK-----VFGVASQ 350
 Db 300 -----RTLN-----LTEVDGKFP-WAFDSAPVGFPDF-GKCDWHMIRISKTP 339
 QY 351 RNLDSSTTRAHEAKVDTTAGRTPKLSLEISTSDDFDQ--NQPT-----KFT 396
 Db 340 NNTSSGDPMSRSVSQVNTVQGVFPHLSIQ-----FDEVFNHPTGDYTGTEWISQPT 392
 QY 397 PVGIGVDNEAEFQOWSLPDYSGQFTHNMMLAPAVNFFEQQLLFFRSQPSGGRS--N 454

Db 393 PPG-----TDINLWEIPDYSSLSQAANLAPPVPPPGFGEALVYFVSAPPGPNRSAPN 446
QY 455 GVLDCLPQEWYQHFYQESAPAQTVLVRYVNPDTGKVLFEAKLHKLGFMTIANN--GD 512
Db 447 DV-PCLLPQEXYTHFVSEQPTMGDAALLHYVDPDTNRNLGEFKLYPGGYLTCVPNGVGA 505
QY 513 SPITVPPNGYFRFESWVNPFFYTLAPMGTGN-----GRRRI 547
Db 506 GFQQLPLNGVFLFVSWSRFYQLKEVGTASTARSRLGVRI 546

Search completed: June 1, 2004, 13:45:54
Job time : 49.0765 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 13.0564 Seconds
(without alignments)
2166.837 Million cell updates/sec

Title: US-09-926-799-6
Perfect score: 2896
Sequence: 1 MKMASNDAPSDGAGLVP.....VNPFYTLAPMGNGRRRIQ 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1211	41.8	530	US-08-486-049-3	Sequence 3, Appli
2	285.5	9.9	622	US-09-590-020-6	Sequence 6, Appli
3	281	9.7	623	US-09-590-020-4	Sequence 4, Appli
4	279.5	9.7	626	US-09-590-020-7	Sequence 7, Appli
5	279.5	9.7	668	US-09-617-594A-4	Sequence 4, Appli
6	276	9.5	623	US-09-590-020-2	Sequence 2, Appli
7	262	9.0	669	US-09-617-594A-2	Sequence 2, Appli
8	128.5	4.4	2137	US-09-134-001C-4463	Sequence 4463, Ap
9	121	4.2	776	US-08-961-083-86	Sequence 86, Appl
10	121	4.2	776	US-09-536-784-86	Sequence 86, Appl
11	119	4.1	2318	US-09-091-219-24	Sequence 24, Appl
12	119	4.1	2318	US-09-660-541-24	Sequence 24, Appl
13	117	4.0	2206	US-07-852-260-2	Sequence 2, Appli
14	117	4.0	2206	US-08-461-503-2	Sequence 2, Appli
15	117	4.0	2206	US-08-465-250-2	Sequence 2, Appli
16	114.5	4.0	3892	US-09-328-352-5503	Sequence 5503, Ap
17	108.5	3.7	883	US-09-489-039A-13542	Sequence 13542, A
18	107.5	3.7	1043	US-08-928-361B-30	Sequence 30, Appl
19	107.5	3.7	1721	US-08-700-651-5	Sequence 5, Appli
20	107.5	3.7	1721	US-08-928-361B-6	Sequence 6, Appli
21	107.5	3.7	1721	US-09-588-995A-6	Sequence 6, Appli
22	104	3.6	1257	US-08-750-152A-2	Sequence 2, Appli
23	102	3.5	540	US-08-956-171E-5225	Sequence 5225, Ap
24	102	3.5	540	US-08-956-171E-5242	Sequence 5242, Ap
25	102	3.5	1560	US-09-264-512B-2	Sequence 2, Appli
26	101	3.5	1447	US-09-041-886-25	Sequence 25, Appl
27	101	3.5	1447	PCT-US94-05277-2	Sequence 2, Appli

28	101	3.5	2232	3	US-09-091-219-25	Sequence 25, Appl
29	101	3.5	2232	4	US-09-660-541-25	Sequence 25, Appl
30	101	3.5	2247	3	US-09-091-219-2	Sequence 2, Appli
31	101	3.5	2247	4	US-09-660-541-2	Sequence 2, Appli
32	100.5	3.5	3788	4	US-09-336-447A-76	Sequence 76, Appl
33	99.5	3.4	968	4	US-09-228-986-76	Sequence 76, Appl
34	97	3.3	1042	3	US-08-928-361B-11	Sequence 11, Appl
35	97	3.3	1042	4	US-09-588-995A-11	Sequence 11, Appl
36	97	3.3	1837	3	US-08-928-361B-5	Sequence 5, Appli
37	97	3.3	1837	4	US-09-588-995A-5	Sequence 5, Appli
38	96.5	3.3	1056	4	US-09-079-030-217	Sequence 217, App
39	96.5	3.3	4536	4	US-09-180-432B-27	Sequence 27, Appl
40	96.5	3.3	4536	4	US-09-079-030-1	Sequence 1, Appli
41	95.5	3.3	401	4	US-08-858-207A-344	Sequence 344, App
42	95.5	3.3	447	4	US-09-627-376-10	Sequence 10, Appl
43	95	3.3	501	2	US-08-448-603A-2	Sequence 2, Appli
44	95	3.3	501	3	US-09-134-075-2	Sequence 2, Appli
45	95	3.3	501	4	US-09-492-739-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-486-049-3
; Sequence 3, Application US/08486049
; Patent No. 6572862
; GENERAL INFORMATION:
; APPLICANT: Bates, Mary K
; APPLICANT: Jiang, Xi
; APPLICANT: Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Ave., N.W.
; CITY: Washington, D.C.
; STATE:
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,049
; FILING DATE: June 7, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311,023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-486-049-3

Query Match 41.8%; Score 1211; DB 4; Length 530;
Best Local Similarity 45.4%; Pred. NO. 4.8e-119;
Matches 260; Conservative 74; Mismatches 161; Indels 78; Gaps 11;

Qy 1 MKMASNDAPSDGAG---LVPEIN-NEAMALDPVAGAAIAAPLTGQNIIDPHMINNF 56
DB 1 MMASKDATSSVDGASGAGQLVPEYNASDPLAMDPEVAGSTAVATAGQVNPIDPWIINN 60

```

186 LRANNSGDDVETVSC--RVLTSPSDFSNFLVPPVPTVESKTKPFTPLTITSEMNSRFP 243
250 L-INPYANTNMSGGIVETKPGDFKHLKPEG-----SMLTHGSVPDLIP 298
244 ----VPIESLHTSPFENIVVQCONGRVTLDBELMGTTLQSPQICAFRGVLTSTRSRASD 299
299 KSSSLWIGNRHSWSDITDFIR-----PFVQANRHFDFNOE 334
300 QADATATPRLFNYYVHVDNLNGTVPDPAED--IPG-PLGTDPDR-----GKVFVGAS 349
335 TAGWSTPRFRPITITVSESNMKSGLIGVATDVIIVGIDGWPDTTIPBQLTPAGIYSITA 394
350 QRLNDGTTR----AHEAKVDTTAGRPTPKLGSLEISTDSDDFDQNOPTKFTPVGIGVDNE 405
395 SNGTDTITTAAGYDAAETIVNTTNFKSMYICGSLQRAWDDKKISN---TAFITTAVRKGN 451
406 AE 407
452 IE 453

RESULT 3
US-09-590-020-4
; Sequence 4, Application US/095900020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilinis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 623
; TYPE: PR
; ORGANISM: Feline calicivirus
; US-09-590-020-4

Query Match 9.7%; Score 281; DB 4; Length 623;
Best Local Similarity 26.9%; Pred. No. 1.8e-20;
Matches 90; Conservative 52; Mismatches 132; Indels 60; Gaps 12;

QY 67 SPRNSPGEVLLNLELGPENIPYLAHARMYNGYAGGFEVQVVLGNAFTAGKIIFAAIIPP 126
DB 134 STSETQKILFKQSLGPLNLYSLAKLYVAVSGSVFRFSISGSGVFGKLAIVVPP 193
QY 127 NF-PIDNLSAAQITWCPHVIVDVRQLEPNLPMDDVRNPNFHYNOGSDRLRLIAMLYTP 185
DB 194 GIEPVQSTSMLOQ---PHVLFDARQVEPVIFAIPDLRSNLYHLMSDITDT-TSLVIMVYND 249
QY 186 LRANNSGDDVETVSC--RVLTSPSDFSNFLVPPVPTVESKTKPFTPLTITSEMNSRFP 243
DB 250 L-INPYANTNMSGGIVETKPGDFKHLKPEG-----SMLTHGSVPDLIP 298
QY 244 ----VPIESLHTSPFENIVVQCONGRVTLDBELMGTTLQSPQICAFRGVLTSTRSRASD 299
DB 299 KSSSLWIGNRHSWSDITDFIR-----PFVQANRHFDFNOE 334
QY 300 QADATATPRLFNYYVHVDNLNGTVPDPAED--IPG-PLGTDPDR-----GKVFVGAS 349
DB 335 TAGWSTPRFRPITITVSESNMKSGLIGVATDVIIVGIDGWPDTTIPBQLTPAGIYSITA 394
QY 350 QRLNDGTTR----AHEAKVDTTAGRPTPKLGSLEISTDSDDFDQNOPTKFTPVGIGVDNE 405
DB 395 SNGTDTITTAAGYDAAETIVNTTNFKSMYICGSLQRAWDDKKISN---TAFITTAVRKGN 451

us-09-926-799-6-rai

57 VOAPGGEFTVSPRNSPGEVLLNLELGPENIPYLAHARMYNGYAGGFEVQVVLGNAFTA 116
61 VOAPGGEFTISPNNTPGDVLPSLUGPHLNFLLHLSQMYNGWGNRVRIMLAGNAFTA 120
117 GKIIIFAAIPPNPFIIDNLSSAQITWCPHVIVDVRQLEPNLPMDDVRNPNFHYNOGSDRL 176
121 GKIIIVSCIPPFGSHNLITQAATLPHVIADVRTLDPIEVLPLEDRVNLFFHNDRNQTM 180
177 RLIAMLYTPLRANNSGDDVETVSCRVLTSPSDFSNFLVPPVPTVESKTKPFTPLTITISE 236
181 RLVCMLYTPLRAGGTGDSFVAVGRVMTCPSPDFNLEFLVPVTEQKTRPFTLNLPLSS 240
237 MNSRPPVPIESLHTSPFENIVVQCONGRVTLDBELMGTTLQSPQICAFRGVLTSTRSR 296
241 LNSRAPLPISSMGISPDNVQSVQFQNGRCTLDGLRVGTTPVSLSHVAKIRG-----TSN 295
297 ASDQADATATPRLFNYYVHVDNLNGTVPDPAEDIPGLGTDPDRG-----KYFVGAS 349
296 GT-----VINLTDLGTPFPFEG-PAPIGFPDLOGCDWHIINNTQFHS 339
350 QRLNDSTTRAHEAKVDTTAGRPTPKLGSLEISTDSDDFDQNOPTKFTPVGIGVDN----- 404
340 QTOYD-----VDITPDTFPHLGSIQAN-----GIGSGNVVGYL 373
405 -----EABFQOQMSLDYSGQFTHNNMNLAPAVAPNFPGEQLLFFRSQLPSSGGRSN 454
374 SWISPSHPSGSQVLDWKIPNYGSSITEATHLAPSVVPPGFEVLVFPFMSKMPGPGAYN- 432
455 GVIDCLVPQWQVHFOYESAPAQTOVALRVYVNPDTGKVLFEAKLHKLGMFTIANGDS- 513
433 --LPCLLPQBYISHLASEQAPTVEAALLHYVDFDGTGRNLGEFKAYPDGFLICVFNAGSS 490
514 -PIVPPNGYFRPESWNPVFTLAPMGTGNRR 545
491 GPQQLPINGVFVFSWSRFLKPVGTASSAR 523

RESULT 2
US-09-590-020-6
; Sequence 6, Application US/095900020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilinis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 622
; TYPE: PR
; ORGANISM: Feline calicivirus
; US-09-590-020-6

Query Match 9.9%; Score 285.5; DB 4; Length 622;
Best Local Similarity 26.5%; Pred. No. 6e-21;
Matches 96; Conservative 52; Mismatches 151; Indels 63; Gaps 13;

QY 67 SPRNSPGEVLLNLELGPENIPYLAHARMYNGYAGGFEVQVVLGNAFTAGKIIFAAIIPP 126
DB 134 STSETQKILFKQSLGPLNLYSLAKLYVAVSGSVFRFSISGSGVFGKLAIVVPP 193
QY 127 NF-PIDNLSAAQITWCPHVIVDVRQLEPNLPMDDVRNPNFHYNOGSDRLRLIAMLYTP 185
DB 194 GIEPVQSTSMLOQ---PHVLFDARQVEPVIFAIPDLRSNLYHLMSDITDT-TSLVIMVYND 249
```

RESULT 4

US-09-590-020-7

; Sequence 7, Application US/09590020

; Patent No. 6355246

; GENERAL INFORMATION:

; APPLICANT: Kruger, John M

; APPLICANT: Maes, Roger K

; APPLICANT: Vilnis, Alvars

; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE

; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF

; FILE REFERENCE: MSU 4.1-446

; CURRENT APPLICATION NUMBER: US/09/590,020

; CURRENT FILING DATE: 2000-06-08

; PRIOR APPLICATION NUMBER: 60/138,484

; PRIOR FILING DATE: 1999-06-10

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 7

; LENGTH: 626

; TYPE: PRT

; ORGANISM: Feline calicivirus

US-09-590-020-7

Query Match 9.7%; Score 279.5; DB 4; Length 626;

Best Local Similarity 22.1%; Pred. No. 2.6e-20;

Matches 122; Conservative 70; Mismatches 175; Indels 185; Gaps 22;

QY 67 SPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGGFEVQVVLGNAFTAGKIIFAAIPP 126
DB 134 STSETQGGKILFKQSLGPLNPLYLHAKLYVWSSGIEVRFSGSGVFGGKLAIVVPP 193
QY 127 NFPIDNLSAAQITMCPHVIVDVRQLEPNLMPDVRNNFFHYNQSGDSRLRLIAMLYTPL 186
DB 194 G--VDPVQSTMLQYPHVLFDAQVEPVIFCLPDLRLSTLYHLMSDITD--TSLVIMYNDL 250
QY 187 RANNGDDVFTVSC--RVLTRSPDFSNFLVPP-----TVESK----- 223
DB 251 -INPYANDTSSGCIIVTETKPGDPFKPHLLKPGSMLTHGSPVPSDLIPKSSSLMIGNRH 309
QY 224 -----TKPFTL-----PI-LTISEMSNSRFPVPIESLHTSP 253
DB 310 WSDITDFIIRPFVQANRHFQNBQTAGSTFRFRPISVITTEQNGAKLGIGV-----A 363
QY 254 TENIVVQCGQ--RVTLGELMGTTQLLPSCICAFRGVLTSTRSTSRASDAQDTATPRLFNY 311
DB 364 TDVIVPGIPDGWPDITIPGEL-----IPAGDYAITNGTGN----- 398
QY 312 YHHVQDNLNGTYPDPAEDI PGPLGTPDPRGKVGVSQRNLDSTTRAHEAKVDTTAGRF 371
DB 399 -----DITATGYDTADIIK---NNTNFRGMYICGSLQRAWG-----DKKISNTAFIT 443
QY 372 TPKLGSLEISTDSD-----FDQOPTKFTPVGIGVDNEAEFQWMSLPDYSQGFTH 422
DB 444 TATL-----DGDNNKINPCNTIDQSKIIVFQDAHVKKKAQTSDDTLALLGYTG----- 492
QY 423 NNNLAPAVANPFGEQLL-----FFRSQLPSSGGRSNGVLDCLVPQEWVQH---FYQES 473
DB 493 -----IGEAIGSDRDRVVRISTLTETGARGG-----NHPFIYKNS 528
QY 474 AP-----AQTVALLRYVNP-----DTGKVLPEAKLHLKGFMT 506
DB 529 IKLGVVIRSIDVNSQILHTSRQLSNHLYLLPPDSFAVYRIIDNSGWNFDIGDSDGFSF 588
QY 507 IANNGDSPITVP 518
DB 589 VGVSGFGKLEFP 600

RESULT 5

US-09-617-594A-4

; Sequence 4, Application US/09617594A

; Patent No. 6541458

; GENERAL INFORMATION:

; APPLICANT: Kruger, John M

; APPLICANT: Maes, Roger K

; APPLICANT: Vilnis, Alvars

; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE

; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF

; FILE REFERENCE: MSU 4.1-446

; CURRENT APPLICATION NUMBER: US/09/590,020

; CURRENT FILING DATE: 2000-06-08

; PRIOR APPLICATION NUMBER: 60/138,484

; PRIOR FILING DATE: 1999-06-10

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 623

; TYPE: PRT

; ORGANISM: Feline calicivirus

; APPLICANT: Audonnet, et al.

; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V

; FILE REFERENCE: 454313-3151.1

; CURRENT APPLICATION NUMBER: US/09/617,594A

; CURRENT FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: 60/193,332

; PRIOR FILING DATE: 2000-03-30

; PRIOR APPLICATION NUMBER: France 00 01761

; PRIOR FILING DATE: 2000-02-11

; PRIOR APPLICATION NUMBER: France 99 09421

; PRIOR FILING DATE: 1999-07-16

; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 4

; LENGTH: 668

; TYPE: PRT

; ORGANISM: Feline calicivirus

US-09-617-594A-4

Query Match 9.7%; Score 279.5; DB 4; Length 668;
Best Local Similarity 24.7%; Pred. No. 2.9e-20;
Matches 94; Conservative 45; Mismatches 108; Indels 133; Gaps 16;

QY 67 SPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGGFEVQVVLGNAFTAGKIIFAAIPP 126
DB 179 STSETQGGKILFKQSLGPLNPLYLHAKLYVWSSGIEVRFSGSGVFGGKLAIVVPP 238
QY 127 NFPIDNLSAAQITMCPHVIVDVRQLEPNLMPDVRNNFFHYNQSGDSRLRLIAMLYTPL 186
DB 239 G--IDPVQSTMLQYPHVLFDAQVEPVIFIPDLRNSLYHLMSDITD--TSLVIMYNDL 295
QY 187 -----RANNGDDVFTVSCRVLTSPDFSNFLVPPVPTVESKTKPFTLPILTISEMSNSR 241
DB 296 INPYANDTSSGCIIVTETKPGDPFKPHLLKPP----- 328
QY 242 FVPVIESLHTSFTENIVVQCQNGRVTLGELMGTTQLLPSCICAFRGVLTSTRSTSRASDOA 301
DB 329 -----GSMLTGHSI--PSDLIPKSSSLMIG----- 351
QY 302 DTATPRLFNYHHVQDNLNGTYPDPAEDI PGPLGTPDPRGKVGVSQRNLDSTTRAHE 361
DB 352 -----NRHW-----SDI-----TDFVIKPFVFOANRHFDFN----- 377
QY 362 AKVDTTAG-----RFTP-----KGLSLEISTDS-----DDP--DQOQPTKFTPVG 399
DB 378 --OETAGWSTPRFRPITITVSEKGSKLG--IGVATDSIVPGIPDGWPDITTIPEKLTPAG 433
QY 400 -IGVDNEAEFQWMSLPDYSG 418
DB 434 DYAITNGGNNDITTAADYDG 453

RESULT 6

US-09-590-020-2

; Sequence 2, Application US/09590020

; Patent No. 6355246

; GENERAL INFORMATION:

; APPLICANT: Kruger, John M

; APPLICANT: Maes, Roger K

; APPLICANT: Vilnis, Alvars

; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE

; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF

; FILE REFERENCE: MSU 4.1-446

; CURRENT APPLICATION NUMBER: US/09/590,020

; CURRENT FILING DATE: 2000-06-08

; PRIOR APPLICATION NUMBER: 60/138,484

; PRIOR FILING DATE: 1999-06-10

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 623

; TYPE: PRT

; ORGANISM: Feline calicivirus

QY 456 VLDCLVPOEWHQYQESAPAOQVALVRYVNPDTGKVLFEAKHLKLGFWMTIANGDSPI 515
Db 525 -----LQOVQGTPEYTESA-----VAKVRYVDANTGKIIPPK-----TIAGEVDATV 568
QY 516 TV 517
Db 569 NI 570

RESULT 9
US-08-961-083-86
; Sequence 86, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 776 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-083-86

Query Match 4.2%; Score 121; DB 3; Length 776;
Best Local Similarity 20.5%; Pred. No. 0.0023;
Matches 110; Conservative 75; Mismatches 233; Indels 118; Gaps 25;

QY 31 PVAGAAIAAPLTGQCOQNIIDPWIMNFVQAPGGEFTVSPRNSPCEVLLNLELGPINPYLA 90
Db 58 PLAFASDSNPDPVSNVDKLSYN--NQPANRWNTWNETNP-EASVGLFGD----- 107
QY 91 HLARMYNGVAGGFEVQVVLGNAFTAGKIIFRAIPNPFDNLNSAAQITMCP----- 142
Db 108 -----SGILSKRSVDNLSVGFHEDHG-----VGVPKSYVIEYVGVKTPTAPKNPSFVGN 157
QY 143 --HVIDVVRQLEPV-NLPPDV-----RNFFHYNQSDSLRL-----IA 180
Db 158 EDHVFVDSANWKEVNTLKPAPQLKAGEMNHFGDKVETAVIRVMVKADNKRGT SITEVQ 217
QY 181 MLYTPLRANNSGDVFTVSCRVLTRPSDPFSNFLVPPTVESKTRPFTLPILTISEMSNS 240
Db 218 IFAKQVAAKQGTQIRQVQDKLANFNPLTDIYL--ESVDGK-----VPATVAS-VSN 269
QY 241 RFPVPIESLHTSPTENIVVQCONGRVTLDELMTGTTQLLPSQICAFRGVLTSTRASDQ 300

Db 270 GLATVVPVSREGEPRVRIAKAEN-----GDILGBYRL-----HFTKDKSLLSHKPVAANKQ 320
QY 301 A-----DTATPRLENYIMHVQLDNLNGTYPD--PAEDIPGFLGTPDFRGKVFGVA 348
Db 321 ARLLQVQGALELPTKVPVYFTGKCGYETKOLTVHEEVPANLT-KAGQFTVGRVLG-- 377
QY 349 SORNLDSTTRAHEAKVDTTAGFTPKLGLSLEISTSDDFDQNOPTKFTPVGIGVNEAEF 408
Db 378 --SNL-----VAEITVRVTDKLG--ETLSDPNYDENSNAQFASATNDIDKNSHD 423
QY 409 QQMSLPDYSGQFTHN---MNLAPAVAPNPPGQQLFERSQLPSSSGRSNG-VLDCILVPOE 464
Db 424 RVDYLD--GDHSENRRTNWSPTSSNPEVSAGVIP-----RENGKIVERTVTQ 472
QY 465 WYQHIFYQESAPAOQVALVRYVNPDTGKVLFEAKHLKLGFWMTIANGDSPIVPPN 520
Db 473 KVQFPADSGTDAPSKVLERYVGPB-----FVPTYSNYQ--AYDADHPNPNPEN 521

RESULT 10
US-09-536-784-86
; Sequence 86, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 776 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 86:
US-09-536-784-86

Query Match 4.2%; Score 121; DB 4; Length 776;
Best Local Similarity 20.5%; Pred. No. 0.0023;
Matches 110; Conservative 75; Mismatches 233; Indels 118; Gaps 25;

QY 31 PVAGAAIAAPLTGQCOQNIIDPWIMNFVQAPGGEFTVSPRNSPCEVLLNLELGPINPYLA 90
Db 58 PLAFASDSNPDPVSNVDKLSYN--NQPANRWNTWNETNP-EASVGLFGD----- 107
QY 91 HLARMYNGVAGGFEVQVVLGNAFTAGKIIFRAIPNPFDNLNSAAQITMCP----- 142

Db 108 -----SGILSKRSVDNLISVGFHEDHG-----VGVEPKSYVIEYYVGKVTPTAPKNPSFVGN 157
Qy 143 -----HVIVDVROLEPV-NLMPDV-----RNNFFHYNQSDSRLRL-----IA 180
Db 158 EDHVFNDSANWKPVTNLKAPQALKAGEMWHFSDFKVEYAVIRVMKADNKRGTSTIEVQ 217
Qy 181 MLYTPLRANNSGDVFTVSCRYLTPRSPDFSNFLVPPVTESKTKPFTPLPILITISEMSN 240
Db 218 IFAKQVAAKQGTQIQVDGKDLANFNPLDLYL--ESVDGK-----VPAVTAS-VSN 269
Qy 241 RFPVPIESLHTSTENIVVQCONGRVTLDELGMGTQLLPSQICAFRGVLTRSTSRASQADATPRFLN 310
Db 270 GLATVVPSVREGEPVRIAKAEN-----GDILGEYRL-----HFTKDKLSLHKFVAAVKQ 320
Qy 301 A-----DTATPRLFNYWHVQLDNLNGTPYD-PAEDI PGPLGTPDPFRGKVFVGA 348
Db 321 ARLLQVQGALELPTKVPVFTGKGVEYTKDLIVEMEVEVPAENLT-KAQQFTVRGRVLG-- 377
Qy 349 SORNLDSTTRAHBAKVDYTAGRTPKLGSLEISTDSDPDQNPQKFTFPGVIGVDNEAEF 408
Db 378 --SNL-----VAEITVRVTDKLG--ETLSDNPNYDENSNOAFASATNDIDKNSHD 423
Qy 409 QWLSLPDYSQFTHN--WNLAFAVAPNPPGQLLFFRSQPSGGRSNG-VLDCILVPOE 464
Db 424 RVDYLAND--GDHSENRRTWNSPTPSNPEVSAGVIF-----RENGKIVERTVIQG 472
Qy 465 WQHFYQESAPAQTOVALRYVNPDTGKVLFEAKLHKLGMFTIANNQSDSPITVPPN 520
Db 473 KVQFFADSGTDASKVLERYVGP-----FEVPTVYSNYQ--AYDADHPFNPNPEN 521

RESULT 11
US-09-091-219-24
; Sequence 24, Application US/09091219
; Patent No. 6171592
; GENERAL INFORMATION:
; APPLICANT: STUDDERT, Michael J.
; APPLICANT: CRABB, Brendan S.
; APPLICANT: FENG, Li
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
; CURRENT FILING DATE: 1998-10-05
; PRIOR APPLICATION NUMBER: US/09/091,219
; CURRENT FILING DATE: 1998-10-05
; PRIOR APPLICATION NUMBER: PCT/AU96/00815
; EARLIER FILING DATE: 1996-12-18
; EARLIER FILING DATE: 1995-12-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 2318
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
US-09-091-219-24

Query Match 4.1%; Score 119; DB 3; Length 2318;
Best Local Similarity 19.5%; Pred. No. 0.024;
Matches 106; Conservative 60; Mismatches 221; Indels 156; Gaps 24;

Qy 79 LELGPEINPYLAHLARMYNGYAGGFEVQVVLGNAFTAGKIIFAAIPNFPIDNLSAAQI 138
Db 367 LELPTDHKGVYGLTDSYAYMNGWDVEVTAAGNFGCGCLLVAMVPELYSIQRELYQL 426
Qy 139 TWCPHVIVDVRQLEPVLNMPDVRNFFHYNQSDSR-LRLIAMLTYPLRANNSG----- 192
Db 427 TLFPHQFINPRNTMHAITVPFVGVN--RYDQYKVKHKEPWLVMVVAELTVNTEGAPOIK 484
Qy 193 -----DDVFTVSCR-----VLTTP-----SPDFSFNFLVP----- 217
Db 485 VYANIAPTNNHVAGPEPSKEGIFPVACSDGYGGLVTTDPKADPVYGVKFNPPRNLPCR 544
Qy 218 -----PTVESKT-----KPFTLPILTISEMSNSRFPVIESLH 250
Db 545 FTNLLDVAEACPTFLRFEQGVYVTKTDSRVLQAFDMS-LAAQKMSNT-FLAGLAQY 602
Qy 251 TSPTENIVQCONGRVTLDELGMGTQLLPSQICAFRGVLTRSTSRASQADATPRFLN 310

Db 545 FTNLLDVAEACPTFLRFEQGVYVTKTDSRVLQAFDMS-LAAQKMSNT-FLAGLAQY 602
Qy 251 TSPTENIVQCONGRVTLDELGMGTQLLPSQICAFRGVLTRSTSRASQADATPRFLN 310
Db 603 TQ-----YSGTINLHFMFTPTDADKARWVAY-----APPGMEPPKTPAAAA 644
Qy 311 YYWHVOLD-NLN-----GTPYDPAEDIPGLGTDPFRGKVFVASORNLDSTT---RAHE 361
Db 645 HCIHAEMWDTGLNSKFTFSIPY-----LSAADYAVTASGVAETTNVQGVWVCLFQITH 695
Qy 362 AKYDVTAGRTPKLGSLEISTDSDPDQNPQKFTFPGVIGVDNEAEFQWMSLPDYSQFT 421
Db 696 GKADGDA-----LWLASAGKDFELRLP-----VDRATTS-----AGE-- 730
Qy 422 HNNMLAPAVAPNPPGQLLFFRSQPSGGRSNGVLDCLV---PQWVQHFYQESAPAQ 478
Db 731 -SADPVTVTVENGGTQQRQHTDVS-----FIMDRFVKVTFQNOINILDMQIPSH 784
Qy 479 QVALRYVNPDTGKVLFEAKLHKLGMFTIANNQSDSPITVPPNG--YFRFESWNP-FYTL 535
Db 785 LVG-----ALLRASTYFSDLEIAVKHEGDLTWVPNGAPEKALDNTTNTAYHK 833
Qy 536 APM 538
Db 834 APL 836

RESULT 12
US-09-660-541-24
; Sequence 24, Application US/09660541
; Patent No. 6531136
; GENERAL INFORMATION:
; APPLICANT: STUDDERT, Michael J.
; APPLICANT: CRABB, Brendan S.
; APPLICANT: FENG, Li
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
; CURRENT FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: US/09/660,541
; CURRENT FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/091,219
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: AU PN7201
; PRIOR FILING DATE: EARLIER FILING DATE: 1995-12-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 2318
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
US-09-660-541-24

Query Match 4.1%; Score 119; DB 4; Length 2318;
Best Local Similarity 19.5%; Pred. No. 0.024;
Matches 106; Conservative 60; Mismatches 221; Indels 156; Gaps 24;

Qy 79 LELGPEINPYLAHLARMYNGYAGGFEVQVVLGNAFTAGKIIFAAIPNFPIDNLSAAQI 138
Db 367 LELPTDHKGVYGLTDSYAYMNGWDVEVTAAGNFGCGCLLVAMVPELYSIQRELYQL 426
Qy 139 TWCPHVIVDVRQLEPVLNMPDVRNFFHYNQSDSR-LRLIAMLTYPLRANNSG----- 192
Db 427 TLFPHQFINPRNTMHAITVPFVGVN--RYDQYKVKHKEPWLVMVVAELTVNTEGAPOIK 484
Qy 193 -----DDVFTVSCR-----VLTTP-----SPDFSFNFLVP----- 217
Db 485 VYANIAPTNNHVAGPEPSKEGIFPVACSDGYGGLVTTDPKADPVYGVKFNPPRNLPCR 544
Qy 218 -----PTVESKT-----KPFTLPILTISEMSNSRFPVIESLH 250
Db 545 FTNLLDVAEACPTFLRFEQGVYVTKTDSRVLQAFDMS-LAAQKMSNT-FLAGLAQY 602
Qy 251 TSPTENIVQCONGRVTLDELGMGTQLLPSQICAFRGVLTRSTSRASQADATPRFLN 310


```

; REFERENCE/DOCKET NUMBER: 36607-D-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEFAX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-503-2

Query Match 4.0%; Score 117; DB 2; Length 2206;
Best Local Similarity 18.6%; Pred. No. 0.036;
Matches 121; Conservative 71; Mismatches 197; Indels 260; Gaps 28;

QY 11 SNDGAAGLVPEINNEAMALDPA-----GAATAAPLT----- 42
Db 270 TNSATIVLPYVN--ALAIIDSMVKHNNWGIALPLSLDFAQDSSVEIPITVIAPMCSE 327
QY 43 --GQONIIDP-----WIMNFFVQAPGG--EFTVSPR-NSPGEV----- 75
Db 328 ENGLRNVTAPEQGLPVLTGSGNQYLTSNHSQPCAIPEFDVTPPIDIPGEVKNMELA 387
QY 76 -----LNLLELGPENPYLAH-----LARMY 96
Db 388 EIDTMIPLNLESTKNTMDMYRVTLSADLSQPLCLSLSPAFDPRLSHTMLGEVLNY 447
QY 97 NGVAGGFEVQVVLGNAGTAKIIFAAIPPNFPIDNLSAAQITMCPHVIVDVRLQEPVNL 156
Db 448 THWAGSLKFTFLFCGSMMATGKILVAYAPPGAQ--PPTSKEAMLGTHVINDLGLQSSCTM 506
QY 157 PMPDVRNFFHYNGS--DS-----RLRLIAMLYTPLRANSGDDVFTVSCRVL 204
QY 507 VVPWISN--VTYRQTQDSFTEGGYISMFYQTRIVVPLSTPKSMMLG---FVSACN--- 558
Db 507 VVPWISN--VTYRQTQDSFTEGGYISMFYQTRIVVPLSTPKSMMLG---FVSACN--- 558
QY 205 RPSDFPSFNF-----VPTVESKTKPFTLPILTIS-----EMSNSRFPVPPIES 248
Db 559 ----DFSVLLRDTTHISQSALPOGIEDLTSEVAGALTLSPKQDLSLPTKASGPAHS 614
QY 249 LHTSPTENIVQCGRVTLDGELMGTTQLLPQICAFRGVL--TRSTSRASDAQDATPR 307
Db 615 KEV-PALTAV-----ETGATNPLAPSDTVQTRHVQVRRSRSESTIESFFA--- 658
QY 308 LFNYWHVQDLNLTGTPYDPAEDIPGLTDPFGKVGFGVASQRLNLSSTTRAHEAKVDIT 367
Db 659 -----RGACVAIIIEVDNEQPTTRAQ----- 678
QY 368 AGRFPTKLGSLSTSDSDPDQONQPKFTPVGIGVDNEAEFQWLSLPDYSGQFTH---NM 424
Db 679 -----KLFAMWRITYKDTVQLRKLEF-----FTYSRFD 708
QY 425 NLAPAVAPNFPGEQLLFFRSQPSGGRSNGVLDCL-----VPOEWQVHFYQES---- 473
Db 709 EFTFVVTANFTN-----ANNGHALNQVYQIMYIPPCAPTPKSWDDYTWQTSNPS 758
QY 474 -----APAQTVQVALRVNPDGKVLFEAKLHLKFWTIANN--GDS 513
Db 759 IFETYGAAPARISVPYVGLANAYSH---FYDGFAKVPKLTQDANDQIGDS 804

RESULT 15
US-08-465-250-2
; Sequence 2, Application US/08465250
; Patent No. 6136570
; GENERAL INFORMATION:
; APPLICANT: Racanietello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
;
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release 1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,250
FILING DATE: 6-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 36607-E-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2206 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-465-250-2

Query Match 4.0%; Score 117; DB 3; Length 2206;
Best Local Similarity 18.6%; Pred. No. 0.036;
Matches 121; Conservative 71; Mismatches 197; Indels 260; Gaps 28;

QY 11 SNDGAAGLVPEINNEAMALDPA-----GAATAAPLT----- 42
Db 270 TNSATIVLPYVN--ALAIIDSMVKHNNWGIALPLSLDFAQDSSVEIPITVIAPMCSE 327
QY 43 --GQONIIDP-----WIMNFFVQAPGG--EFTVSPR-NSPGEV----- 75
Db 328 ENGLRNVTAPEQGLPVLTGSGNQYLTSNHSQPCAIPEFDVTPPIDIPGEVKNMELA 387
QY 76 -----LNLLELGPENPYLAH-----LARMY 96
Db 388 EIDTMIPLNLESTKNTMDMYRVTLSADLSQPLCLSLSPAFDPRLSHTMLGEVLNY 447
QY 97 NGVAGGFEVQVVLGNAGTAKIIFAAIPPNFPIDNLSAAQITMCPHVIVDVRLQEPVNL 156
Db 448 THWAGSLKFTFLFCGSMMATGKILVAYAPPGAQ--PPTSKEAMLGTHVINDLGLQSSCTM 506
QY 157 PMPDVRNFFHYNGS--DS-----RLRLIAMLYTPLRANSGDDVFTVSCRVL 204
QY 507 VVPWISN--VTYRQTQDSFTEGGYISMFYQTRIVVPLSTPKSMMLG---FVSACN--- 558
Db 507 VVPWISN--VTYRQTQDSFTEGGYISMFYQTRIVVPLSTPKSMMLG---FVSACN--- 558
QY 205 RPSDFPSFNF-----VPTVESKTKPFTLPILTIS-----EMSNSRFPVPPIES 248
Db 559 ----DFSVLLRDTTHISQSALPOGIEDLTSEVAGALTLSPKQDLSLPTKASGPAHS 614
QY 249 LHTSPTENIVQCGRVTLDGELMGTTQLLPQICAFRGVL--TRSTSRASDAQDATPR 307
Db 615 KEV-PALTAV-----ETGATNPLAPSDTVQTRHVQVRRSRSESTIESFFA--- 658
QY 308 LFNYWHVQDLNLTGTPYDPAEDIPGLTDPFGKVGFGVASQRLNLSSTTRAHEAKVDIT 367
Db 659 -----RGACVAIIIEVDNEQPTTRAQ----- 678
QY 368 AGRFPTKLGSLSTSDSDPDQONQPKFTPVGIGVDNEAEFQWLSLPDYSGQFTH---NM 424
Db 679 -----KLFAMWRITYKDTVQLRKLEF-----FTYSRFD 708
QY 425 NLAPAVAPNFPGEQLLFFRSQPSGGRSNGVLDCL-----VPOEWQVHFYQES---- 473
Db 709 EFTFVVTANFTN-----ANNGHALNQVYQIMYIPPCAPTPKSWDDYTWQTSNPS 758
QY 474 -----APAQTVQVALRVNPDGKVLFEAKLHLKFWTIANN--GDS 513
Db 759 IFETYGAAPARISVPYVGLANAYSH---FYDGFAKVPKLTQDANDQIGDS 804
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 35.0315 Seconds
(without alignments)
4368.312 Million cell updates/sec

Title: US-09-926-799-6
Perfect score: 2896
Sequence: 1 MKMASNDAPSDGAGLVP.....VNPFYTLAPMTGTGRRRIQ 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1211	41.8	530	14	US-10-314-739-3
2	279.5	9.7	668	14	US-10-209-507-4
3	279.5	9.5	547	12	US-10-670-695-4
4	275.5	9.5	671	12	US-10-670-695-2
5	262	9.0	669	14	US-10-209-507-2
6	126	4.4	1965	15	US-10-369-493-2646
7	121.5	4.2	522	14	US-10-172-502-17
8	121.5	4.2	1831	12	US-10-282-122A-71033
9	121	4.2	338	10	US-09-867-932A-2
10	121	4.2	776	9	US-09-765-272-86
11	121	4.2	2233	10	US-09-769-787-2
12	116.5	4.0	1147	15	US-10-327-481A-38
13	112.5	3.9	932	14	US-10-174-677-88
14	112	3.9	709	12	US-10-231-956A-287
15	111.5	3.9	485	14	US-10-172-502-18
					Sequence 3, Appli
					Sequence 4, Appli
					Sequence 4, Appli
					Sequence 2, Appli
					Sequence 2, Appli
					Sequence 2646, Ap
					Sequence 17, Appl
					Sequence 71033, A
					Sequence 2, Appli
					Sequence 86, Appl
					Sequence 2, Appli
					Sequence 38, Appl
					Sequence 88, Appl
					Sequence 287, App
					Sequence 18, Appl

16	111.5	3.9	2271	12	US-10-282-122A-43924	Sequence 43924, A
17	111.5	3.9	2283	14	US-10-172-502-4	Sequence 4, Appli
18	111	3.8	5175	15	US-10-120-801-74	Sequence 74, Appl
19	111	3.8	5175	15	US-10-369-493-6859	Sequence 6859, Ap
20	111	3.8	5175	15	US-10-369-493-6861	Sequence 6861, Ap
21	111	3.8	5198	12	US-10-346-863-34	Sequence 34, Appl
22	111	3.8	5198	15	US-10-120-801-75	Sequence 75, Appl
23	111	3.8	5198	15	US-10-369-493-6858	Sequence 6858, Ap
24	111	3.8	5198	15	US-10-369-493-6860	Sequence 6860, Ap
25	110.5	3.8	5877	14	US-10-142-515-11	Sequence 11, Appl
26	110.5	3.8	5935	14	US-10-243-243A-8	Sequence 8, Appli
27	109.5	3.8	2344	9	US-09-815-242-12713	Sequence 12713, A
28	109	3.8	6310	12	US-10-282-122A-67793	Sequence 67793, A
29	109	3.8	19662	15	US-10-084-846A-6	Sequence 6, Appli
30	108.5	3.7	997	15	US-10-369-493-3593	Sequence 3593, Ap
31	108.5	3.7	2005	9	US-09-735-367B-3	Sequence 3, Appli
32	108.5	3.7	2063	9	US-09-735-367B-2	Sequence 2, Appli
33	107.5	3.7	487	15	US-10-369-493-22178	Sequence 22178, A
34	107	3.7	1686	15	US-10-386-414-2	Sequence 2, Appli
35	107	3.7	2478	9	US-09-815-242-5816	Sequence 5816, Ap
36	107	3.7	2478	9	US-09-815-242-12967	Sequence 12967, A
37	106.5	3.7	1794	10	US-09-965-738-299	Sequence 299, App
38	106.5	3.7	1799	10	US-09-965-738-149	Sequence 149, App
39	106.5	3.7	1821	10	US-09-965-738-82	Sequence 82, Appl
40	106.5	3.7	2234	12	US-10-612-090-20	Sequence 20, Appl
41	106.5	3.7	2358	12	US-10-282-122A-45763	Sequence 45763, A
42	106.5	3.7	11721	10	US-09-965-738-162	Sequence 162, App
43	106	3.7	920	15	US-10-320-797-3323	Sequence 3323, Ap
44	105.5	3.6	435	16	US-10-389-566-1432	Sequence 1432, Ap
45	105	3.6	1255	12	US-10-312-352-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-10-314-739-3
; Sequence 3, Application US/10314739
; Publication No. US20030129588A1
; GENERAL INFORMATION:
; APPLICANT: Estes, Mary K
; Jiang, Xi
; Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; Characterize No. US20030129588A1walk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Ave., N.W.
; CITY: Washington, D.C.
; STATE: <Unknown>
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/10/314, 739
; APPLICATION NUMBER: US/10/314, 739
; FILING DATE: 09-Dec-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/486, 049
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311.023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX: <Unknown>

```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
; US-10-209-507-4

Query Match          9.7%; Score 279.5; DB 14; Length 668;
Best Local Similarity 24.7%; Pred. No. 1.5e-17;
Matches 94; Conservative 45; Mismatches 108; Indels 133; Gaps 16;

QY 67 SPNRSPEVLLNLELGPINPYLAHARMYAGGVQVVLGNAGTAKIIFAAIIPP 126
DB 179 STSETQKILFKQSLGSLPLNPLYLTHLAKLYAVAGSIEVRSISGSGVFGKLAIVVPP 238
QY 127 NPFIDNLAAQITMCPHVIVDVRQLEPNLMPDVRNNFFHYNOGDSRLRLIAMLYTPL 186
DB 239 G--IDPVQSTSMQYPHVLFDARQVEPVIFIPDLRNSLYHLMSDITD-TSLIVIMIYNDL 295
QY 187 -----RANNSGDDVFTVSCRVLTRPSDPFSFNLVPTVESKTKPFTLPILTISEMSNR 241
DB 296 INFYANDSNSGCIIVTVE-----TKPGDPFKFLLKPP----- 328
QY 242 FVPFIESLHTSPENIVVQCGRVTLDGELMGTTLQLLPSQICAFRGVLTSTRASDOA 301
DB 329 -----GSMUTHGSI--PSDLIFKSSSLWIG----- 351
QY 302 DTATPRLFNYYHVQLDNLNGTYPDPAEDIPGLPTDPRGRKGVGVASORNLDSITRAHE 361
DB 352 -----NRHM-----SDI-----TDFVIKPFVFOANRHFDFN----- 377
QY 362 AKVDTTAG-----RPTP-----KLGSLRISTDS-----DDF-DQNOPTKFTPVG 399
DB 378 ---QETAGWSTPRPRPTITVSEKSGSKUG-IGVATDSIVPGIPDGPMDPTTPEKLTAG 433
QY 400 -IGVDNEAEFQOWSLPDYSG 418
DB 434 DYAITNGGNDITTAADYDG 453

RESULT 3
US-10-670-695-4
; Sequence 4, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; TITLE OF INVENTION: STATUS OF AN ANIMAL
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Feline calicivirus
; US-10-670-695-4

Query Match          9.5%; Score 275.5; DB 12; Length 547;
Best Local Similarity 22.3%; Pred. No. 2.7e-17;
Matches 123; Conservative 70; Mismatches 174; Indels 185; Gaps 23;

QY 67 SPNRSPEVLLNLELGPINPYLAHARMYAGGVQVVLGNAGTAKIIFAAIIPP 126
DB 55 STSETQKILFKQSLGSLPLNPLYLTHLAKLYAVAGSIEVRSISGSGVFGKLAIVVPP 114
QY 127 NPFIDNLAAQITMCPHVIVDVRQLEPNLMPDVRNNFFHYNOGDSRLRLIAMLYTPL 186
DB 239 G--IDPVQSTSMQYPHVLFDARQVEPVIFIPDLRNSLYHLMSDITD-TSLIVIMIYNDL 295
QY 187 -----RANNSGDDVFTVSCRVLTRPSDPFSFNLVPTVESKTKPFTLPILTISEMSNR 241
DB 296 INFYANDSNSGCIIVTVE-----TKPGDPFKFLLKPP----- 328
QY 242 FVPFIESLHTSPENIVVQCGRVTLDGELMGTTLQLLPSQICAFRGVLTSTRASDOA 301
DB 329 -----GSMUTHGSI--PSDLIFKSSSLWIG----- 351
QY 302 DTATPRLFNYYHVQLDNLNGTYPDPAEDIPGLPTDPRGRKGVGVASORNLDSITRAHE 361
DB 352 -----NRHM-----SDI-----TDFVIKPFVFOANRHFDFN----- 377
QY 362 AKVDTTAG-----RPTP-----KLGSLRISTDS-----DDF-DQNOPTKFTPVG 399
DB 378 ---QETAGWSTPRPRPTITVSEKSGSKUG-IGVATDSIVPGIPDGPMDPTTPEKLTAG 433
QY 400 -IGVDNEAEFQOWSLPDYSG 418
DB 434 DYAITNGGNDITTAADYDG 453

US-10-209-507-4
; Sequence 4, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26

```


Db 115 G--VDPQSTSLMQYPHVLFDAQVEPVIFCPLDKSTLYHLMSTDT--TSLVIMVYNDL 171
Qy 187 RANNSGDDVFTVSC--RVLTSPSPFQSFNLFVPP-----TVES-----KT-----224
Db 172 -INPANDANSSGCVTVETKPGDFKPHLLKPPGSMTHGSIPLDLPKTSLSLWGNRY 230
Qy 225 -----KPTTL-----PI-LTISEMNSRPPVPPIESLHTSP 253
Db 231 WSDITDFVIRPVFOANRHFDFNQETAGNSTRFRPISVTITEQNGAKLIGV-----A 284
Qy 254 TENIVVQONG--RVTLDELGMGTQLLPSQICAFRGVLTSTRSASDAQADTATPRLFNY 311
Db 285 TDYIVPGIPDGWPDFTIPGEL-----IPAGDYAITNGTN-----319
Qy 312 YHVHQLDNLNGTPYPAEDI PGPLGTDPFRGVGVFVASORNLDSSTTRAHEAKVDTTAGRF 371
Db 320 -----DITATGYDTADIK--NNTNFRGMVYICSLQRAWG-----DKISNTAFIT 364
Qy 372 TPGLSLEISTDSD--FDQOQTKFTPVGIGVDNEABFQWMSLPDYSQGFTH 422
Db 365 TATL-----DGDNNKINPCNTIDQSKI VVQDNHVGKKAQTSDDTLALLGYTG----413
Qy 423 NNNLAPAVAPNPFGEQLL-----FFRSQLPSSGSRNGVLDCLVPQEWVQH--FYQBS 473
Db 414 -----IGEQAIGSDRDRVVRISTLPETGARG--NHPIFYKNS 449
Qy 474 AP-----AQTOVALVRVNP-----DTGKVLPEAKLHLKLGFT 506
Db 450 IKLGVVIRSIDVFNSQLHTSRQLSLNHYLLPPDSFAVYRIIDNSNGSWFIDIGSDGFSF 509
Qy 507 IANNGDSPITVP 518
Db 510 VGVSGFGKLEFP 521

RESULT 4

US-10-670-695-2
; Sequence 2, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; TITLE OF INVENTION: STATUS OF AN ANIMAL
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-670-695-2

Query Match 9.5%; Score 275.5; DB 12; Length 671;
Best Local Similarity 22.3%; Pred. No. 3.8e-17;
Matches 123; Conservative 70; Mismatches 174; Indels 185; Gaps 23;
Qy 67 SPNSPGEVLLNLELGPENPYLAHARMYNGVAGFEVQVVLGNAFTAGKIIFAADPP 126
Db 179 STSETQKILFKQSLGPLNLYLHLAKLYVAMSGSIEVRSISGSGVFGGKLAIVVPP 238
Qy 127 NFPIDNLSAAQITMCCHVIVDROLEPVNPLMPDVNRNFFHYNQSDSLRLIAMLTYPL 186
Db 239 G--VDPQSTSLMQYPHVLFDAQVEPVIFCPLDKSTLYHLMSTDT--TSLVIMVYNDL 295
Qy 187 RANNSGDDVFTVSC--RVLTSPSPFQSFNLFVPP-----TVES-----KT-----224

Db 296 -INPANDANSSGCVTVETKPGDFKPHLLKPPGSMTHGSIPLDLPKTSLSLWGNRY 354
Qy 225 -----KPTTL-----PI-LTISEMNSRPPVPPIESLHTSP 253
Db 355 WSDITDFVIRPVFOANRHFDFNQETAGNSTRFRPISVTITEQNGAKLIGV-----A 408
Qy 254 TENIVVQONG--RVTLDELGMGTQLLPSQICAFRGVLTSTRSASDAQADTATPRLFNY 311
Db 409 TDYIVPGIPDGWPDFTIPGEL-----IPAGDYAITNGTN-----443
Qy 312 YHVHQLDNLNGTPYPAEDI PGPLGTDPFRGVGVFVASORNLDSSTTRAHEAKVDTTAGRF 371
Db 444 -----DITATGYDTADIK--NNTNFRGMVYICSLQRAWG-----DKISNTAFIT 488
Qy 372 TPGLSLEISTDSD--FDQOQTKFTPVGIGVDNEABFQWMSLPDYSQGFTH 422
Db 489 TATL-----DGDNNKINPCNTIDQSKI VVQDNHVGKKAQTSDDTLALLGYTG----537
Qy 423 NNNLAPAVAPNPFGEQLL-----FFRSQLPSSGSRNGVLDCLVPQEWVQH--FYQBS 473
Db 538 -----IGEQAIGSDRDRVVRISTLPETGARG--NHPIFYKNS 573
Qy 474 AP-----AQTOVALVRVNP-----DTGKVLPEAKLHLKLGFT 506
Db 574 IKLGVVIRSIDVFNSQLHTSRQLSLNHYLLPPDSFAVYRIIDNSNGSWFIDIGSDGFSF 633
Qy 507 IANNGDSPITVP 518
Db 634 VGVSGFGKLEFP 645

RESULT 5

US-10-209-507-2
; Sequence 2, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-2

Query Match 9.0%; Score 262; DB 14; Length 669;
Best Local Similarity 22.3%; Pred. No. 7.6e-16;
Matches 110; Conservative 69; Mismatches 166; Indels 148; Gaps 20;
Qy 30 DPVAGAAATAAPLTGQONTIDPMMNMFVQAPGGEFTV-----SPRNSPGEVLLNLELGP 83
Db 146 EPSAQMATADATKGVSDSEW--ESF-----FSFHTSVNWSTSETQKILFKQSLP 196
Qy 84 EINYLAHLARMYNGVAGFEVQVVLGNAFTAGKIIFAIPNFPIDNLSAAQITMCPH 143
Db 197 LLNPLYLHLKLYVAMSGSDVVRFSISGSGVFGGKLAIVVPPG--VDPVQSTSLMQYPH 254
Qy 144 VLVVROLEPVNPLMPDVNRNFFHYNQSDSLRLIAMLTYPL-----RANNSGDDVFTV 198
Db 255 VLFDAQVEPVIFCPLDKSTLYHLMSTDT--TSLVIMVYNDLNPANDNSNGSICVTV 313

```

199 QY   SCRVLTRPSDFSNFLVPP-----TVESKTKPTLL 229
      :|:|||||: |||
314 Db E---TRGPDFKFHLKLPGGSMLTCHGISPSDLIPKSSSLWIGNRWYSDIITDFVIRPEVF 369
      :|:|||||: |||
230 QY   -----PI-LTISEMSNRFPVPPIESLHTSPENIVVOQONG-RV 266
      :|:|||||: |||
370 Db QANRHFDNQETAGCWSTPRFRPITIITISENGSKLTGV-----ATDYIVPGIDGWPD 423
      :|:|||||: |||
267 QY TLDGELMGTTOALLSQICAFRGVLTRTSRASDOADATPRLENYVHHVQLDNLANGTPVD 326
      :|:|||||: |||
424 Db TTIGS-----ELITFA-----GDYSIINGNDIA-----TANAYD 453
      :|:|||||: |||
327 QY PAEDIPGLGPDPDRFGKVFGVASORNULDDTTTRAHEAKVDTAGRFTPKLSLEISTDSOD 386
      :|:|||||: |||
454 Db SADVI---TNNTNFRGMVICGALORAW-----DKKISSTA-FITTAKBGNWLKPSMT 503
      :|:|||||: |||
387 QY PQONQTKFTFVGIGVDNEAEFAQOWSLPDYSGOPTHNMNLAPAVAPNFPGCQLLFRR---- 443
      :|:|||||: |||
504 Db IDMTKIANVQDTHVRDRVQTSDDLLAILGYTG-----IGEQAIGNSNRDS 547
      :|:|||||: |||
444 QY ----SOLPSSGOR 452
      :|:|||||: |||
548 Db VVRISMILPETGAR 560
      :|:|||||: |||

```

RESULT 6
 US-10-369-493-2646
 ; Sequence 2646, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPANSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 2646
 ; LENGTH: 1965
 ; TYPE: PRT
 ; ORGANISM: *Synechocystis* sp.
 ; US-10-369-493-2646

	Query Match	4.4%;	Score 126;	DB 15;	Length 1965;	
	Best Local Similarity	20.9%;	Pred. No. 0.061;			
	Matches 133;	Conservative	76;	Mismatches 241;	Indels 186;	Gaps 29;
Qy	3	MASNDAAEPCNGAAGLVPEINNEAMALDPVAGAAIAAPLTGQQNIIDPIMMNFVOAPGG	62			
Dd	969	IATDDADVSAGITFGKLP- <u>GND-----</u> AGSFTINPTIGKVTLIDD---PFQTSQSY	1016	:	:	:
Qy	63	EFTVGR--NSPGEVILNL-----ELGPENPYLAHARMWGVAGGFEVQVVIAGNA	113	:	:	:
Dd	1017	KFTVTASDGVNLTAEQLVNLTIIDLDEIAPTITSGDTAKALAENGANQIVIEVIATDDA	1076	:	:	:
Qy	114	F7AGKIIPAAIP-----FNFPIDNLSSAQITMCBHVIVDVROLE	152	:	:	:
Dd	1077	DVSAGITFGLKGNDAQSFTINPTIGKVTLIDDPDFETQSKYKFTVTASDGVNLTAEQL-	1135	:	:	:
Qy	153	PVNLPMPDYRN-----NPHYNQGSRSRLRTAMLATPLRANNNG	192	:	:	:
Dd	1136	-VNLIITLIDEIAPTITSGDTAKAITENSGANIYYQATADDNADIAGVTFGLKPGDDA	1194	:	:	:
Qy	193	DD--VFVTSRVLTRPSDF----SFNFLVPPTVESKKFPTLPILATISEMSNSRFVPI	246	:	:	:
Dd	1195	DSPFINAIQTQVLLNDPDFEQSSVKFTV--TASDGVNPTEQVNVNLTIIDLDEIAPTI	1252	:	:	:

```

QY      247 ESLTSP--TEN-----IVVQCNGRVTL--    268
           |||::|||
Db     1253 TSGNTAKAIANGSANGNQVIYQTATDDADISAEVTFLKLGDDATSFTSTTKGVTLIN 1312
           |||::|||
QY     269 ---DGELMGTQLLPSCICAFRGVLTRTSRASDAQDATPRLFNYWHVDNLN-----   320
           |||::|||
Db    1313 NPDEFESQSXYKF-----TVTASDGVNPATEKIYT---LAINNLDIEGPT 1353
           |||::|||
QY     321 ---NGTPYPDAEDIPGF-----LGTPD---FRGKVFGVASQRNLDSTTRAHEAKVDTTA 368
           |||::|||
Db    1354 ITSGDTATPIDENSCANGLIYQVIATDDADVSAGITFGL-----KPGNDA 1398
           |||::|||
QY     369 GRFT--PKLGSLEISTSDDEDFONOPTKTFTPVGIGVDNEAEFQQWSLP----DYSGQFTH 422
           |||::|||
Db    1399 GSFITNPITGEVTL-IDDPDFETQSYSKYFTVTASDGVNPATEKIVLTINDLDELGPNI 1457
           |||::|||
QY     423 NMNLAPAVANPFPGBLLAFFRSQLPSSGR--SNVGDLCLVPQEWQHVFQESAPAQTQV 480
           |||::|||
Db    1458 SDNIATAINENSANCQIIY---QVTADDGDIDISAGVTFLKPGDNANOFSID--PTTGQV 1512
           |||::|||
QY     481 ALVRYNPDGTGKVLPEAKHLKGEMTIANNGDSPIT 516
           |||::|||
Db    1513 TLDD--NPD-----FETQ-SSYKFVTASDGVNPAT 1540
           |||::|||

RESULT 7
US-10-172-502-17
; Sequence 17, Application US/10172502
; Publication NO. US20030105833A1
; GENERAL INFORMATION: Timothy et al.
; APPLICANT: POSTER, CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES...
; TITLE OF INVENTION: P07263US01/BAS
; FILE REFERENCE: P07263US01/BAS
; CURRENT APPLICATION NUMBER: US/10/172,502
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/298,098
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 522
; TYPE: PR
; ORGANISM: Staphylococcus epidermidis
US-10-172-502-17
```

[illegible]

QY 508 ANNGDSPITV 517
Db 462 AGEVDGTNI 471

RESULT 8

US-10-282-122A-71033
; Sequence 71033, Application US/10282122A
; Publication No. US20040029129A1

GENERAL INFORMATION:

; APPLICANT: Wang, Liangau
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 71033

; LENGTH: 1831

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

; FEATURE:

; NAME/KEY: MISC FEATURE

; LOCATION: (1801)..(1801)

; OTHER INFORMATION: X=any amino acid

; FEATURE:

; NAME/KEY: MISC FEATURE

; LOCATION: (1805)..(1805)

; OTHER INFORMATION: X=any amino acid

; FEATURE:

; NAME/KEY: MISC FEATURE

; LOCATION: (1808)..(1808)

; OTHER INFORMATION: X=any amino acid

; FEATURE:

; NAME/KEY: MISC FEATURE

; LOCATION: (1809)..(1809)

; OTHER INFORMATION: X=any amino acid

; FEATURE:

; NAME/KEY: MISC FEATURE

; LOCATION: (1825)..(1825)

; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1828)..(1828)
; OTHER INFORMATION: X=any amino acid
US-10-282-122A-71033

Query Match 4.2%; Score 121.5; DB 12; Length 1831;

Best Local Similarity 23.5%; Pred. No. 0.15;

Matches 87; Conservative 39; Mismatches 131; Indels 113; Gaps 21;

QY 219 TVESKTKPFTLPILITISEMSNSRFPVPIESLHT-----SPT-----E 255

Db 236 TSEARTSTNQISNLTSTSTNOSSPTSPANLRTFRFTVLNTMAAPTSTTTTSLTGN 295

QY 256 NIVVQCQN-----GRVTLDELMTGTTQLLPISOICAPRGVLTSTSTASQADATPR 307

Db 296 SVVVKDNFNEHMLNLSGSATYDPK-TGIATLTDPAY-SQKGAISLNLRLDSN-----R 346

QY 308 LPNYYVHVOLDNLNGTPYDP-----AEDI-----PGPLGTPDFRGKVFVASORNLDSTT 357

Db 347 SPFRFGKVNLSN-RYEGYSPDGVAGDGGIGFAFSPGLQIGKEGAAVIGGLNN----- 400

QY 358 RAHEAKVDTTAGRTPK-----LGSLEISTSDDDFDQONQTKTPVGI 400

Db 401 -AFGFKLDITYHNTSTPRSDAKAKADPRNVGGGAGFAGF-VSTDRNGMATTESTAALKLV 458

QY 401 -GVDNEAEFOQWSLPDYSQ-----FTNMNLAPAVAPNFGCEQLLFFRSQLP 447

Db 459 OPTDN-SFQDFVI-DYNGDTKVMVTYAGQTFTRNLT---DWIKNSGG--TTFSLSWTA 510

QY 448 SSGRSNGVLDCLVPQEWQHFYQBSAPAQTOVALRVYVNPDTGKVLPEAKHLKLGFMFI 507

Db 511 STGGAKN-----LQOVQFGTEYTESA-----VAKRVYVDANTGKDIIPPK-----TI 553

QY 508 ANNGDSPITV 517

Db 554 AGEVDGTNI 563

RESULT 9

US-09-867-932A-2

; Sequence 2, Application US/09867932A

; Publication No. US20030049825A1

GENERAL INFORMATION:

; APPLICANT: Lin, Han-You

; APPLICANT: Kuo, Tsun-Yung

; APPLICANT: Huang, Hsiao-I

; APPLICANT: Yang, Huey-Lang

; TITLE OF INVENTION: NERVOUS NECROSIS VIRUS PROTEIN

; FILE REFERENCE: 08919-023001

; CURRENT APPLICATION NUMBER: US/09/867,932A

; CURRENT FILING DATE: 2001-05-30

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 338

; TYPE: PRT

; ORGANISM: Nervous necrosis virus

US-09-867-932A-2

Query Match 4.2%; Score 121; DB 10; Length 338;

Best Local Similarity 24.6%; Pred. No. 0.011;

Matches 82; Conservative 44; Mismatches 141; Indels 66; Gaps 17;

QY 8 AAPSDNGAAGLVP--EINNEAMALDPVAGAAIAAPLTG---QQNIIDPWIMNFVQA--P 60

Db 10 AKPATTKAANQPORRRANRRNRNRTDAPVSKASTVTGFGRCNTDNLHSLGMSRISQAVLP 69

QY 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGVAG--GFEVQVLGNAFTAG 117

Db 70 AGTGT-----DGYVVVDATIVPDLPLRLGHAARIPQRYAVETLEFEIQPMCPAN--TGG 121

118 KIIIPAIIPNPIDN--LSAAQITMCPHVIIVDRQLEPVNLPMPDVRNFFHYNQSDS 174
 122 GYVAGFLPD--PTDNDHTFGALQATR-CAVAVAKWWSRTVR--PQVTRTLLWTSGKEQ 175
 175 RLRLIAMLYTLPLRANSGDDVFTVS--CRVLTRSPSPDFSNFLVPPVTSKTKFTLPIL 232
 176 RLTSRGLILLCVGNNT--DVVNVSVLCRWSVR-----LSVPSLETPEET-TAPIM 223
 233 TISEMSNRFPPIESLHTSPENTIVVQCONGRVTLGELMGTQTQLLPQSCAIFRGLVTR 292
 224 TQGSYLN-----DSLTNSDKSILLGSTPLDIAPDGAVQLDRLSIDYSLGTGDVDR 276
 293 STSRASDAQATATPRLFNYHYHVLQDNLN-GTP 324
 277 AV-----YHLLKKFAGNAGTP 292

RESULT 10
 US-09-765-272-86
 ; Sequence 86, Application US/09765272
 ; Patent No. US20020061545A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Choi et. al.
 ; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
 ; NUMBER OF SEQUENCES: 452
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/765,272
 ; FILING DATE: 22-Jan-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/961,083
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brookes, A. Anders
 ; REGISTRATION NUMBER: 36,373
 ; REFERENCE/DOCKET NUMBER: PB340P2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 86:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 776 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 86:
 US-09-765-272-86

Query Match 4.2%; Score 121; DB 9; Length 776;
 Best Local Similarity 20.5%; Pred. No. 0.042;
 Matches 110; Conservative 75; Mismatches 233; Indels 118; Gaps 25;

31 PVAGAAIAAPLTGQQNIIDPWIMNMFVQAPGGEFTVSPNSPGEVLLNLELGPINPYLA 90
 58 PLAFASDSNPSDPVSNVDKLIISYN--NQPANRWNTWRTNP-EASVGLFGD----- 107
 91 HLARMYNGVAGGEVQVILAGNAFTAGKIIFAAIPNPFPIDNLSAAQITMCP----- 142
 108 -----SGILSKESVDNLVSGFHEDHG-----VGVPKSYVIEYVGVKVTPTAPKNPSFVGN 157

143 --HVIIVDRQLEPV-NLPMPOV-----RNNFFHYNQSDSRLRL-----IA 180
 158 EDHVFNDSANWKPVTNLKAPQLKAGENMHFSEDKVETAVRIRMVKADNRKGTISITEVQ 217
 181 MLYTFLRANSGDDVFTVSCRVLTRSPSPDFSNFLVPPVTSKTKFTLPILITISEMSNS 240
 218 IFAKQVAAAKQCGQTRIQVDGKDLANFNPDLDIYLL--ESVDGK-----VPAVTAS-VSN 269
 241 RFPVPFESLHTSPTENIVVQCONGRVTLGELMGTQTQLLPQSCAIFRGLVTRSTRASDQ 300
 270 GLATVVPVSREGEPVRIKAEN-----GDILGEVRL--HFTKDKSLLSHKPVAAVKQ 320
 301 A-----DTATPRLFNYHYHVLQDNLNNGTTPD-PAEDIPGLTPTDPGRGVFGVA 348
 321 ARLLQVGQALELPTKVPVYFTKGQYETKDLTVEEVEPAENLT-KAGQFTVRGVGLG-- 377
 349 SQRNLDSTTRAHEAKYDTAGRTPKLGSLEISTDSDDDFDQONQPTKFTPVGIGVDNEAEF 408
 378 --SNL-----VAEITVRVTDKLG--ETLSNPNYDENSNOAFASATNDIDKNSHD 423
 409 QQWMLPDYSGQFTHN--MNLAPAVAPNPFGEQLLFFRSQLPSSGGRSNG-VLDCLVPQE 464
 424 RVDYLDN--GDHSENRRTWNSPTPSSNPEVSAGVIF-----RENGKIVERTVTOG 472
 465 WVQHVFQESAPATOVALRVYVNPDTGKVLFEAKLHKLGMFTIANNQDSDPITVPPN 520
 473 KVQFFADSGTDPASKLVLERYGVPE-----FEVPTYISNYQ--AYDADHPFPNNPEN 521

RESULT 11
 US-09-769-787-2
 ; Sequence 2, Application US/09769787
 ; Publication No. US200300091577A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Microbial Technics Limited
 ; APPLICANT: Gilbert, Christophe FG
 ; APPLICANT: Hansbro, Philip M
 ; TITLE OF INVENTION: Proteins
 ; FILE REFERENCE: PWC/P21129WO
 ; CURRENT APPLICATION NUMBER: US/09/769,787
 ; CURRENT FILING DATE: 2001-01-26
 ; PRIOR APPLICATION NUMBER: GB 9816337.1
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: US 60/125164
 ; PRIOR FILING DATE: 1999-03-19
 ; NUMBER OF SEQ ID NOS: 388
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 2233
 ; TYPE: PRT
 ; ORGANISM: Streptococcus pneumoniae
 ; US-09-769-787-2

Query Match 4.2%; Score 121; DB 10; Length 2233;
 Best Local Similarity 20.5%; Pred. No. 0.23;
 Matches 110; Conservative 75; Mismatches 233; Indels 118; Gaps 25;

31 PVAGAAIAAPLTGQQNIIDPWIMNMFVQAPGGEFTVSPNSPGEVLLNLELGPINPYLA 90
 1480 PLAFASDSNPSDPVSNVDKLIISYN--NQPANRWNTWRTNP-EASVGLFGD----- 1529
 91 HLARMYNGVAGGEVQVILAGNAFTAGKIIFAAIPNPFPIDNLSAAQITMCP----- 142
 1530 -----SGILSKESVDNLVSGFHEDHG-----VGVPKSYVIEYVGVKVTPTAPKNPSFVGN 1579
 143 --HVIIVDRQLEPV-NLPMPOV-----RNNFFHYNQSDSRLRL-----IA 180
 1580 EDHVFNDSANWKPVTNLKAPQLKAGENMHFSEDKVETAVRIRMVKADNRKGTISITEVQ 1639
 181 MLYTFLRANSGDDVFTVSCRVLTRSPSPDFSNFLVPPVTSKTKFTLPILITISEMSNS 240
 1640 IFAKQVAAAKQCGQTRIQVDGKDLANFNPDLDIYLL--ESVDGK-----VPAVTAS-VSN 1691

```
QY 241 RPPVPIESLHTSPENIVVQCONGRVTLDELMTGTTQLLPSCQICAPRGVLTSTSPASQ 300
Db 1692 GLATVVPVSREGEPRVIAKAEN-----GDILGEYRL-----HFTKDKSLSHKPAVAAVKQ 1742
QY 301 A-----DTATPRLFNFNHVVQDLNLTGTPYD-PAEDIPGPGTGPDPGRKVFVGA 348
Db 1743 ARLLQVGOALELPTKVPVFTGKGVEYTKDLTVHEHEEVAENLT-KAGQFTVRGRVLG-- 1799
QY 349 SORNLDSTTRAHEAKVDTTAGRTPKLGSLEISTSDDDFDQONQPTKFTPVGIGVNEAEF 408
Db 1800 ---SNL-----VAEITVRVTDKLG--ETLSNPNYDENSNOAFASATNDIDKNSHD 1845
QY 409 QWMSLPDYSGQTHN---NNLAPAVAPNFPGLLFFRSQLPSSGGRSNG-VLDCILVPE 464
Db 1846 RVDYLD--GDHSENRRTWNSPTSSNPEVSAGVIF-----RENGKIVERTVTQG 1894
QY 465 WQHFQESAPACTOVALVRYNPDTKVLFKAKLHKLGFMTIANGDSDPITVPPN 520
Db 1895 KVQFFADSGTDAPSKLVLERYVGE-----FEVPTYSNYQ--AYDADHPFNPN 1943

RESULT 12
US-10-327-481A-38
; Sequence 38, Application US/10327481A
; Publication No. US20040001864A1
; GENERAL INFORMATION:
; APPLICANT: King, Andrew M.Q.
; APPLICANT: Burman, Allison J.
; APPLICANT: Audonnet, Jean-Christophe F.
; APPLICANT: Lombard, Michel F.A.
; TITLE OF INVENTION: Vaccine against Foot-and-Mouth Disease
; FILE REFERENCE: 454313-3178
; CURRENT APPLICATION NUMBER: US/10/327,481A
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: PCT/FR01/02042
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: FR 0008437
; PRIOR FILING DATE: 2000-06-29
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 38
; LENGTH: 1147
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
US-10-327-481A-38

Query Match 4.0%; Score 116.5; DB 15; Length 1147;
Best Local Similarity 23.1%; Pred. No. 0.21; Mismatches 25; Indels 27; Gaps 3;
Matches 36; Conservative 25;

QY 70 NSPGEVLLMLELGPENIPYLAHARMYNGYAGGFEVQVVLGNAGTAAFKIIPAAIPPNFP 129
Db 158 DKPFGVLTLELPTDHHGVFGLVDSYAMRWGMDVEVSAGVNGFGCLLVAMVPEWKA 217
QY 130 INLSAAQITMCPHVIVVQLEPNLMPDVNRNFFHYNQGSRSR-LRLIAMLVTPRA 188
Db 218 FDTREKYQLTFEPHOFISPRNTAHITVPYLGVN--RYDQYKHKHPWTLVVMVLSPLTV 275
QY 189 NN-----SGDDVETVSC 200
Db 276 SNTAAPQIKVYANIAFTYVHVHAGELPSKEGIFPVAC 311

RESULT 13
US-10-174-677-88
; Sequence 88, Application US/10174677
; Publication No. US20030190704A1
; GENERAL INFORMATION:
; APPLICANT: Xie, Ting
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ANCHORING STEM CELLS IN A MICROENVIR
; FILE REFERENCE: 40716 (JP-012)
; CURRENT APPLICATION NUMBER: US/10/174,677
; CURRENT FILING DATE: 2002-06-19
```

```
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 88
; LENGTH: 932
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-174-677-88

Query Match 3.9%; Score 112.5; DB 14; Length 932;
Best Local Similarity 20.9%; Pred. No. 0.37;
Matches 112; Conservative 63; Mismatches 175; Indels 185; Gaps 29;

QY 102 GFEVQVVLGNAPT--AGKIIPAAIPPNFPIDNLSAAQITMCPHVIVVQ--LEPVNL 156
Db 12 GFPLLSILLGTTPWEAGRILY-----SVSEETDKGSFVGDTAKDLGLEPREL 59
QY 157 PMPDV-----RNNFFHYNQGSRSRLRLIAMLVTLPLRANNSGDDVFTVSCRVLTRPSPDF 210
Db 60 AERGVRILSRGRTQLFALNQSGS-----LVTAGRIDR--EEICAQSAARCLV----- 104
QY 211 SFNPLVPPPTVESKTK--PFTLPILPTISE-----MSNS-----RPPV----- 244
Db 105 NFNIL-----MEDKNNLYPIDVEIIDIINDNVPFLTEBINVKIMENTAPGVRFPPLSEAGDP 160
QY 245 -----PIESLHTSPENIVVQCONGRVTLDELMTGTTQLLPSCQICAFRGVLTSTSR---- 296
Db 161 DVGTSLSQSYQLSPNRHFLAVQSG---DDETKYPELVL-----ERVLDREBERVHH 209
QY 297 ---ASDQADATPRLFNHVMHVLQDLNLTGTPYDPAEDIPGPGTGPDPGRKVFVGA 352
Db 210 LVLTAASDGGD--PPR--SSTAHIQTVVVDNDHTVPFSLP-----QYQVTVPEN 254
QY 353 LDSTTR---AHEAKVDTTAG-----RFTPKLGLS-----EIST-DSDDFD----- 388
Db 255 VPVGTGRLTVAIDLDGEGVNGEVTVSPRKITPKLPKMFHLNSLTGEISTLEGLDYEETAF 314
QY 389 -----NQPTKFTPVGIGV-----DNEAEFQWLSLPDYSGQTHNNLAPAVAPNPP 435
Db 315 YEMEYQAQDGGSLTKAKVLITVLVDNDNAPEVMTSL---SSSIPEDTPLGTVIA----- 367
QY 436 GSQLLFFRSQLPSSGGRSNGVLDCVLPQEWQHFQESAPACTOVALVRYNPDTKVLF 495
Db 368 ---LFY---LQDRDSGKNGEVTCTIP-----ENLPPKLEKSIDNYRLVTTKNLD 411
QY 496 EAKLHKLK-----FMTIANGDSDPITVPPNGYFRFESWVNP 531
Db 412 RETLSLYNITLKATDGGTGPLLSRETHIFMQVADTNDNPPTPHSHSSYVYIAENNP 466

RESULT 14
US-10-231-956A-287
; Sequence 287, Application US/10231956A
; Publication No. US20040053233A1
; GENERAL INFORMATION:
; APPLICANT: Lorens, James B.
; APPLICANT: Xu, Weiduan
; APPLICANT: Bogenberger, Jakob
; APPLICANT: Holland, Sacha
; APPLICANT: Rigel Pharmaceuticals, Incorporated
; TITLE OF INVENTION: Modulators of Angiogenesis
; FILE REFERENCE: 021044-004100US
; CURRENT APPLICATION NUMBER: US/10/231,956A
; CURRENT FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 287
; LENGTH: 709
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-231-956A-287

Query Match 3.9%; Score 112; DB 12; Length 709;
Best Local Similarity 20.7%; Pred. No. 0.27;
```

Wed Jun 2 09:13:31 2004

249 GVLGETGLNGRAVGIGGLSNAFGFKLDYTHNTSKPNSAAKANADPSNVAGGAGFAGFVTT 308
373 PKGLSLEISTDSDDDFDQN-----QPTKETPVGIGVDNEAEFQWLSLPDYSQ-ETHNMN- 425
309 DSYGVATTVTSSSTADNAAKLVNQPTNNTFQDFDINYNQDITKMTV-KYAGQTWTRNISD 367
426 -LAPAVAPNPFCEQLLFFRSQPLSSGGRSNGVLDCLVPQEWQHFYQESAPAQVQVALVR 484
368 WIAKSTTNFS-----LSMTASTGGATN-----LQQVQFGTPEYTESAVTQ-----VR 410
485 YVNPDTGKVLFEAKHLKGLGFWMTIANNGDSPITV 517
411 YVDVTTGKDIIPK-----TYSGNVDQVVTI 436

Search completed: June 1, 2004, 14:04:37
Job time : 36.0315 secs

Matches 96; Conservative 50; Mismatches 137; Indels 180; Gaps 23;
QY 125 PPNF--PIDNLSAAITWCPHVIVDVRQLE-PVN-----LPMEDVRNFFHYNGS 172
Db 149 PPSFQTPVNTVSSTNLVTPPAVVSQPKLQPTVTSGSLTATSVLPAPNTAT----- 199
QY 173 DSRRLIAMLTYPLRANSGDDVFTVSCR-----VLTTP 206
Db 200 -----VWATTQVP-----SGNQPTISLQPLVILHVPVAVSSQPLLQSHPGTLVTNP 249
QY 207 SPDFSG-NFLVPPTVESKTK-BFTLITISMSNSRFPVIESLHTSPTEN-----I 257
Db 250 SGNVEFISVQSPTVSGLTKNPSLPLNPKPNVPSPSPSQRTNPTASAPLPTL 309
QY 258 VVQC-----QNGRVL-----DGLMGTQLLPSQ-----ICAPRGVLTSTRASDQ 300
Db 310 AVQAVPTAHSIVQATRTSLTVGSPGLYSPSTRNGPIQMKIPIAF-----STSSAAEQ 363
QY 301 ADTATPRLFNYYWHVQLDNLNGTTPYDPAEDIPLGLTDPFRGKVFVGASQRNLDSTTRAH 360
Db 364 NSNTTPIEN-----QTNKI-----IDASVSKAADSTSQCG 395
QY 361 EAK-----VDT-----AGRTPKLSLEISTDSDFDONOPTK-----FTPVG 399
Db 396 KATGSDSSGVIDLTMDDDESGASQDPKLNHTPVTMSSSQPVSRPLQIQPAPPLQPSG 455
QY 400 IGVNNEAEFQWLSLPDYSQGFTHNMNLAFAV-----APNPFGEQLLFFRSQPLSSGGR 452
Db 456 VPTSGSPSTHLLP--TAPTTVNTVTRPVTQVTRLPVPRAPANHVVY-TTLP----- 507
QY 453 SNGVLDCLVPQEWQHFYQESAPAQ-----QVALVRYNP 488
Db 508 -----APPAQAPLRTVMQAPAVRQVNP 530

RESULT 15
US-10-172-502-18
; Sequence 18, Application US/10172502
; Publication No. US20030185833A1
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy et al.
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES...
; FILE REFERENCE: P07263US01/BAS
; CURRENT APPLICATION NUMBER: US/10/172.502
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/298,098
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-172-502-18
Query Match 3.9%; Score 111.5; DB 14; Length 485;
Best Local Similarity 22.6%; Pred. No. 0.16;
Matches 89; Conservative 50; Mismatches 169; Indels 85; Gaps 19;
QY 169 NQGSRLRLIAMLTYPLRANSGDDVFTVSCRVLTRPSPSPFNLVPPPTVESKTKPFT 228
Db 85 NTSSTSDTKSVASTSTEQPTNSTQSTASNTSQSTPS-SVNLNKTSTTSTAP-- 141
QY 229 LPILITISEMSNSRFPVIESLHTSPTENIV-----OCQNGRVLTDGE--LMGTQ 277
Db 142 VKLRTESRLAMSTFASAATT--TAVTANTITVKNKLMKMTYSGNATYDQSTGIVLTQ 199
QY 278 LLPSQICAPRGVLTSTRASDAQDTATPRLFNYYWHVQLDNL-----NGTPYDPAEDIP 332
Db 200 DAYSQ-----KGATITLGRIDSNS-----FHFSKVLNKGNGHNGGDIQGFAPSP 248
QY 333 GPLGTDPFRGKVFVA-----SORNLDSTTRAHAKVDIT-----ACRF-----T 372

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 46.3893 Seconds
(without alignments)
3289.030 Million cell updates/sec

Title: US-09-926-799-7
Perfect score: 2845
Sequence: 1 MKMASNDTPSNDGAAGLVP.....VNQFYSLAPMTGNGRRRVQ 540

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2845	100.0	540	AAB49706	Small rou
2	2269.5	79.8	535	AAB49707	Small rou
3	2018	70.9	548	AAB49705	Small rou
4	2001	70.3	550	AAB49709	Small rou
5	1979	69.6	548	AAB49708	Small rou
6	1933	67.9	542	AAB49704	Small rou
7	1872.5	65.8	539	AAB49704	Small rou
8	1650.5	58.0	541	AAB49710	Small rou
9	1233	43.3	530	AAB50972	Small rou
10	1233	43.3	530	ADC72176	Norwalk v
11	1232	43.3	530	AAB49701	Small rou
12	1225	43.1	545	AAB49700	Small rou
13	1183	41.6	530	AAR57091	Small rou
14	1170.5	41.1	544	AAB49703	Small rou
15	1169.5	41.1	546	AAB49702	Small rou
16	323.5	11.4	579	AAB08143	RHDV caps
17	296.5	10.4	668	AAB67462	Amino aci
18	292.5	10.3	547	AAM50108	Feline ca
19	292.5	10.3	668	AAR10686	Feline ca
20	292.5	10.3	668	AAB04304	Feline ca
21	292.5	10.3	671	AAM50107	Feline ca
22	281	9.9	622	AAB47045	Feline ca
23	281	9.9	623	AAB47044	Feline ca
24	277	9.7	623	AAB47043	Feline ca
25	273	9.6	669	AAB67461	Amino aci

26	186	6.5	40	5	AAU91274	
27	167	5.9	40	5	AAU91273	Norwalk v
28	139	4.9	934	1	AAF20016	Sequence
29	138	4.9	1147	5	ABF76724	Foot and
30	131.5	4.6	338	7	ADC06674	Nervous n
31	123.5	4.3	366	5	ABB05228	Redspotte
32	123	4.3	2209	1	ADP20037	Sequence
33	114.5	4.0	794	7	ADC03508	Raisgonia
34	111	3.9	6304	6	ABU09236	Human neu
35	110.5	3.9	1455	4	AAM79120	Human pro
36	109.5	3.8	689	6	ABR43116	Human IMX
37	109.5	3.8	756	7	ADC31441	Human nov
38	109.5	3.8	2701	5	ABP74125	Human TRI
39	109	3.8	257	6	ABU48168	Protein e
40	109	3.8	992	4	ABB60861	Drosophil
41	108.5	3.8	1125	3	AAB22934	Mouse mic
42	108.5	3.8	1125	3	AAV79637	Microtubu
43	108.5	3.8	1125	5	ABG94497	Protease
44	108.5	3.8	1125	7	ADC18464	Microtubu
45	108.5	3.8	1521	4	ABB52754	Escherich

ALIGNMENTS

RESULT 1
AAB49706
ID AAB49706 standard; protein; 540 AA.
XX
AC AAB49706;
XX
DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 7.
XX
KW Small round structured virus; SRSV; food poisoning.
XX
OS Small round structured virus.
XX
PN WO200079280-A1.
XX
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000WO-JP004095.
XX
PR 22-JUN-1999; 99JP-00175928.
XX
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
(DENK-) DENKA SEIKEN KK.
PA
Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
WPI: 2001-080848/09.
DR N-PSDB; AAF29147.
PT
Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.
XX
Claim 1; Page 54-57; 84pp; Japanese.
PS
This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks
XX
SQ Sequence 540 AA;
Query Match 100.0%; Score 2845; DB 4; Length 540;
Best Local Similarity 100.0%; Pred. No. 6.9e-256;

Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MKMASNDATPSNDGAAGLVPSNNEAMALEPVVVGASLAAPVTGQTNIIIDPWIRTNFVOAP 60
Db	1 MKMASNDATPSNDGAAGLVPSNNEAMALEPVVVGASLAAPVTGQTNIIIDPWIRTNFVOAP 60
QY	61 NGEFTVSPRNSPGEILVNLLELGPENLPYLAHLARMYNGYAGGMEVQVWMLAGNAFTAGKII 120
Db	61 NGEFTVSPRNSPGEILVNLLELGPENLPYLAHLARMYNGYAGGMEVQVWMLAGNAFTAGKII 120
QY	121 FAAVPPYFVENLSPQITMFPFHVIDVRLTPVLLPMPDVRSTLTFHFNQKDEPKRLVA 180
Db	121 FAAVPPYFVENLSPQITMFPFHVIDVRLTPVLLPMPDVRSTLTFHFNQKDEPKRLVA 180
QY	181 MLYTPLRSNGSGDDVFTVSCRILTRPSDFDTYLPVPTVESKTPFTLPVLTGELSNS 240
Db	181 MLYTPLRSNGSGDDVFTVSCRILTRPSDFDTYLPVPTVESKTPFTLPVLTGELSNS 240
QY	241 RFLPSIDEMVTSNSESIVVQPNQGRVTLTGELLGTTQLQACNICSIRGKVTQVPSQHM 300
Db	241 RFLPSIDEMVTSNSESIVVQPNQGRVTLTGELLGTTQLQACNICSIRGKVTQVPSQHM 300
QY	301 WNLITNLNGTQFDDTDDVPAPLGVDPFAGEVFGVLSQRNGESNPANRAHDAVATYSD 360
Db	301 WNLITNLNGTQFDDTDDVPAPLGVDPFAGEVFGVLSQRNGESNPANRAHDAVATYSD 360
QY	361 KYTPKGLVQIGTWTNDVENVQKFTPIGLNEVANGHREFTLPRYSGALTLNNMLAP 420
Db	361 KYTPKGLVQIGTWTNDVENVQKFTPIGLNEVANGHREFTLPRYSGALTLNNMLAP 420
QY	421 AVAPLFGGERLLFFRSYVPLKGGFGNPAIDCSVPQEWVQHFYQESAPSLGDVALVRYVN 480
Db	421 AVAPLFGGERLLFFRSYVPLKGGFGNPAIDCSVPQEWVQHFYQESAPSLGDVALVRYVN 480
QY	481 DTGRVLPFAKLHKGGLTVSSTSTGTVVVPANGYKFDSDWVNFYSLAPMGTGNGRRRVQ 540
Db	481 DTGRVLPFAKLHKGGLTVSSTSTGTVVVPANGYKFDSDWVNFYSLAPMGTGNGRRRVQ 540
RESULT 2	
AAB49707	
ID	AAB49707 standard; protein; 535 AA.
AC	AAB49707;
DT	04-APR-2001 (first entry)
DE	Small round structured virus protein SEQ ID 8.
KW	Small round structured virus; SRSV; food poisoning.
OS	Small round structured virus.
XX	WO200079280-A1.
PD	28-DEC-2000.
PF	22-JUN-2000; 2000WO-JP004095.
PR	22-JUN-1999; 99JP-00175928.
PA	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA	(DENK-) DENKA SEIKEN KK.
PI	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
DR	WPI: 2001-080848/09.
DR	N-PSDB; AAF29148.
PT	Kit for the detection and typing of small round-structured virus (SRSV)
PT	strains for investigation of food poisoning outbreaks, contains
PT	antibodies.
XX	

Claim 1; Page 57-59; 84pp; Japanese.	
This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks	
SQ Sequence 535 AA;	
Query Match 79.8%; Score 2269.5; DB 4; Length 535;	
Best Local Similarity 76.3%; Pred. No. 3.1e-202;	
Matches 413; Conservative 65; Mismatches 56; Indels 7; Gaps 3;	
QY	1 MKMASNDATPSNDGAAGLVPSNNEAMALEPVVVGASLAAPVTGQTNIIIDPWIRTNFVOAP 60
Db	1 MKMASNDATPSNDGAAGLVPSNNEAMALEPVVVGASLAAPVTGQTNIIIDPWIRTNFVOAP 60
QY	61 NGEFTVSPRNSPGEILVNLLELGPENLPYLAHLARMYNGYAGGMEVQVWMLAGNAFTAGKII 120
Db	61 NGEFTVSPRNSPGEILVNLLELGPENLPYLAHLARMYNGYAGGMEVQVWMLAGNAFTAGKII 120
QY	121 FAAVPPYFVENLSPQITMFPFHVIDVRLTPVLLPMPDVRSTLTFHFNQKDEPKRLVA 180
Db	121 FAAVPPYFVENLSPQITMFPFHVIDVRLTPVLLPMPDVRSTLTFHFNQKDEPKRLVA 180
QY	181 MLYTPLRSNGSGDDVFTVSCRILTRPSDFDTYLPVPTVESKTPFTLPVLTGELSNS 240
Db	181 MLYTPLRSNGSGDDVFTVSCRILTRPSDFDTYLPVPTVESKTPFTLPVLTGELSNS 240
QY	241 RFLPSIDEMVTSNSESIVVQPNQGRVTLTGELLGTTQLQACNICSIRGKVTQVPSQHM 300
Db	241 RFLPSIDEMVTSNSESIVVQPNQGRVTLTGELLGTTQLQACNICSIRGKVTQVPSQHM 300
QY	301 WNLITNLNGTQFDDTDDVPAPLGVDPFAGEVFGVLSQRNGESNPANRAHDAVATYSD 360
Db	301 WNLITNLNGTQFDDTDDVPAPLGVDPFAGEVFGVLSQRNGESNPANRAHDAVATYSD 360
QY	361 KYTPKGLVQIGTWTNDVENVQKFTPIGLNEVANGHREFTLPRYSGALTLNNMLA 419
Db	361 KYTPKGLVQIGTWTNDVENVQKFTPIGLNEVANGHREFTLPRYSGALTLNNMLA 419
QY	420 PAVAPLFGGERLLFFRSYVPLKGGFGNPAIDCSVPQEWVQHFYQESAPSLGDVALVRYVN 479
Db	420 PAVAPLFGGERLLFFRSYVPLKGGFGNPAIDCSVPQEWVQHFYQESAPSLGDVALVRYVN 479
QY	479 PAVAPLFGGERLLFFRSYVPLKGGFGNPAIDCSVPQEWVQHFYQESAPSLGDVALVRYVN 539
Db	479 PAVAPLFGGERLLFFRSYVPLKGGFGNPAIDCSVPQEWVQHFYQESAPSLGDVALVRYVN 539
QY	540 Q 540
Db	535 Q 535
RESULT 3	
AAB49705	
ID	AAB49705 standard; protein; 548 AA.
AC	AAB49705;
DT	04-APR-2001 (first entry)
DE	Small round structured virus protein SEQ ID 6.
KW	Small round structured virus; SRSV; food poisoning.
OS	Small round structured virus.
XX	WO200079280-A1.
XX	


```
PD 28-DEC-2000.
XX
XX PF 22-JUN-2000; 2000WO-JP004095.
XX PR 22-JUN-1999; 99JP-00175928.
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI: 2001-080848/09.
XX DR N-PSDB; AAF29146.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
XX Claim 1; Page 52-54; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 548 AA;
SQ
Query Match 70.9%; Score 2018; DB 4; Length 548;
Best Local Similarity 68.3%; Pred. No. 9e-179;
Matches 377; Conservative 71; Mismatches 88; Indels 16; Gaps 4;
QY 1 MKWASNDATPSNDGAAGLVPESSNEMALEPVVGASLAAPVTGQTNIIDPWIRTNFVQAP 60
DB 1 MKWASNDAAAPSNDGAAGLVPEINNEAMALDPVAGAAIAPLTGQONIIDPWIRTNFVQAP 60
QY 61 NGEFTVSPRNSPGEILVNLELGPENLPYLHLARMYNGYAGGMEVQVVLGNAGNAGTAKII 120
DB 61 GGEFTVSPRNSPGEVLNLELGPENLPYLHLARMYNGYAGGFEVQVVLGNAGNAGTAKII 120
QY 121 FAAPPPYFVENVLSQITMFPVHVIDVRLTPLEPVLLPMPDVVSTLTFHFNQKDEPKMRLVA 180
DB 121 FAAPPPYFVENVLSQITMFPVHVIDVRLTPLEPVLLPMPDVVSTLTFHFNQKDEPKMRLVA 180
QY 181 MLYTPLRSNGSGDDVFTVSCRILTRPSPEEDFTYLVPPPTVESKTKPFTLPVLTGLSLS 240
DB 181 MLYTPLRANNSGDDVFTVSCRILTRPSDFSNFLVPPPTVESKTKPFTLPVLTGLSLS 240
QY 241 RPLSIDEMVTSNPESIVVQPNQGRVTLTGELGTTQLQACNICSGIRGKVTGQV--PSEQ 298
DB 241 REPVIESLHTSPTENIVVQCNQGRVTLTGELGTTQLLPSQICAFRGVLTSTSRADQ 300
QY 299 -----HWNLEITNLNGTQEDPTDDVPAPLGVPDFAGEVGVLSQRNRSNAN 348
DB 301 ADTATPRLPNYWHVQLDNLNGTTPYDPAEDIPGLGTPDFRGKVFVASQRN---LDSTT 357
QY 349 RAHDAAVATYSKYTPKGLGVOIGTWNTDVENOQTKFTPIGLNEVANGHREOWTLPRY 408
DB 358 RAHEAKVDVTAAGRTFKPLGSLISITSDSDFDQDQKPTKFTPVGIG-VDNSEAFQOQSLPDY 416
QY 409 SGALTINMNLAPAVAPLPGERLLPFRSVVPLKGGFGNPAIDCSVPQEWQHFYQESAPS 468
DB 417 SQGFTHNMNLAPAVAPNPFGEQLLFRSOLPSSGSGNSGVLDCLVPQEWQHFYQESAPA 476
QY 469 LGDVALRVYNPDGTGRVLFEAKLHKGGLFTVSTSGPVVVPANGYFKFDSWVNFYSIA 528
DB 477 QTOVALRVYNPDGTGRVLFEAKLHKLGFMTIANNGSDSPITVPNGYFRPESWVNFYTLA 536
QY 529 PMGTGNGRRRVQ 540
DB 537 PMGTGNGRRRIQ 548
```

RESULT 4

```
AAB49709
XX
XX ID AAB49709 standard; protein; 550 AA.
XX AC AAB49709;
XX DT 04-APR-2001 (first entry)
XX DE Small round structured virus protein SEQ ID 10.
XX KW Small round structured virus; SRSV; food poisoning.
XX OS Small round structured virus.
XX PN WO200079280-A1.
XX PD 28-DEC-2000.
XX PF 22-JUN-2000; 2000WO-JP004095.
XX PR 22-JUN-1999; 99JP-00175928.
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI: 2001-080848/09.
XX DR N-PSDB; AAF29150.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
XX Claim 1; Page 62-64; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 550 AA;
SQ
Query Match 70.3%; Score 2001; DB 4; Length 550;
Best Local Similarity 67.9%; Pred. No. 3.5e-177;
Matches 377; Conservative 72; Mismatches 86; Indels 20; Gaps 4;
QY 1 MKWASNDATPSNDGAAGLVPESSNEMALEPVVGASLAAPVTGQTNIIDPWIRTNFVQAP 60
DB 1 MKWASNDAAAPSNDGAAGLVPEANDEVMALPEPVVGASIAAPVVGQONIIDPWIRTNFVQAP 60
QY 61 NGEFTVSPRNSPGEILVNLELGPENLPYLHLARMYNGYAGGMEVQVVLGNAGNAGTAKII 120
DB 61 GGEFTVSPRNSPGEVLNLELGPENLPYLHLARMYNGYAGGQVQVVLGNAGNAGTAKII 120
QY 121 FAAPPPYFVENVLSQITMFPVHVIDVRLTPLEPVLLPMPDVVSTLTFHFNQKDEPKMRLVA 180
DB 121 FAAPPPYFVENVLSQITMFPVHVIDVRLTPLEPVLLPMPDVVSTLTFHFNQKDEPKMRLVA 180
QY 181 MLYTPLRSNGSGDDVFTVSCRILTRPSPEEDFTYLVPPPTVESKTKPFTLPVLTGLSLS 240
DB 181 MLYTPLRAN-SEGEDVFTVSCRILTRPAPDFEFLLVPPPTVESKTKPFTLPVLTGLSLS 239
QY 241 RPLSIDEMVTSNPESIVVQPNQGRVTLTGELGTTQLQACNICSGIRGKVTGQV----- 295
DB 240 REPAAIDMLYTDPNESIVVQPNQGRVTLTGELGTTQLQACNICSGIRGKVTGQV----- 299
QY 296 -----SEQHMNLEITNLNGTQEDPTDDVPAPLGVPDFAGEVGVLSQRN-----RGSN 345
```

Db 300 TDSQORARHPLHVQVKLDTQYDPTDDI FAVLGAIDFKTVFGVASQRDVSQQBQGH 359
 Qy 346 PANRAHDAVATYSKYTPKGLGVOIGTWNTNDVENOPTKFTPIGLNEVANGHFEQWTL 405
 Db 360 YATRAHEAHIDTDPKYAPKLGITILIKSGDDFNTNQIRFTIPVGMGD---NNRWQWEL 415
 Qy 406 PRYSALTLNMLNLAFAVAPLFGERLLPFRSVVPLKGGFGNPAIDCSVPQEWVQHFYQES 465
 Db 416 PDYSGLTLNMLNLAFAVSPFGEIRLFRSIVPSAGSGYIDCLIPQEWVQHFYQEA 475
 Qy 466 APLSGDALVRYVNPDTGRVLFPEAKLHKGGLTWSSTSTGPVVPANGYFKFDSWVNOFY 525
 Db 476 AFSQSALVRYVNPDTGRNIFEAKLHREGFLTVANCGNPIVPPNGYFRFEANGNQFY 535
 Qy 526 SLAPMGTGNGRRVQ 540
 Db 536 TLAPMGSGQRRRAQ 550

RESULT 5
 AAU91272
 ID AAU91272 standard; protein; 548 AA.
 XX
 AC AAU91272;
 DT 18-JUN-2002 (first entry)
 XX Norwalk virus associated polynucleotide #1.
 DE Norwalk virus;
 KW Norwalk virus; monoclonal antibody; geno group I; geno group II;
 KW immunological detection; food; viral infection.
 XX Norwalk virus.
 OS
 XX JP2002020399-A.
 XX
 PD 23-JAN-2002.
 PF 10-JUL-2000; 2000JP-00208151.
 XX 10-JUL-2000; 2000JP-00208151.
 PR (OSAP) OSAKA PREFECTURE.
 PA (NISE-) NIPPON SEIBUTSU KAGAKU CENT KK.
 PA (IATR) IATRON LAB INC.
 XX
 DR WPI; 2002-287412/33.
 XX A monoclonal antibody useful in the immunological detection and diagnosis
 PT of Norwalk virus infection.
 XX
 PS Disclosure; Page 12-13; 24pp; Japanese.
 XX
 CC The invention describes a monoclonal antibody recognising Norwalk virus,
 CC a capsid protein of Norwalk virus, or a common antigen epitope on the
 CC capsid protein molecule of geno group I and geno group II. The antibody
 CC is useful for immunological detection and quantitative analysis of
 CC Norwalk virus in foods and the serum of infected patients. This sequence
 CC represents a Norwalk virus associated protein described in the invention
 XX
 XX Sequence 548 AA;
 SQ

Query Match 69.6%; Score 1979; DB 5; Length 548;
 Best Local Similarity 67.9%; Pred. No. 3.9e-175;
 Matches 371; Conservative 71; Mismatches 90; Indels 16; Gaps 4;

Qy 1 MKMASNDATPSNDGAAGLVPSNESAMALEPVPVVGASLAAPVTGQTNIIIDPWIRTFVQAP 60
 Db 1 MKMASRAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQQNIIIDPWIMNPFVQAP 60
 Qy 61 NGFTVSPRNSPGELVNLGELPYNLAHARMYNGYAGGVEVQVVLGNAFTAGKII 120

Db 61 CGEFTVSPRNSPGVLLNLELGPBEINPYLAHARMYNGYAGGFEVQVVLGNAFTAGKII 120
 Qy 121 FNAVPPYFVENLSPSQITMEPHVIIDVRLTEPVLLEMPDVRSTLTFHFNQKDEKMLVA 180
 Db 121 FFAIPNPFIIDNLSSAAQITMCPHVIVDVRQLFEPVNLPMVVRNFFHYNQGSRLRLIA 180
 Qy 181 MLYTLPLRNSGDDVFTVSCRLTRPSPEFDTVLVPTVESKTKPPTLPLVLTGLSLSNS 240
 Db 181 MLYTLPLRANSGDDVFTVSCRLTRPSPDFSFLVPTVESKTKPPTLPLTISEMSNS 240
 Qy 241 RPLSIDEMVTSPNESIVVQONGRVTLDGELGTLTQLOACNICSIRKGVQGV--PSEQ 298
 Db 241 RFPVPTESLHTSPTENIVVQONGRVTLDGELMGTTQLLPRICAFGRVLRSTSRASDQ 300
 Qy 299 -----HWNLEITNLNGTQDPTDDVPAPLGVDPDFAGEVPGVLVSRNRGESNPAN 348
 Db 301 ADIATPRLFNYYHVQLDNLNGTPYDPAEDIPGLGTFDFRGKVGVSQRN---PDSTT 357
 Qy 349 RAHDVAVATYSKYTPKGLGVOIGTWNTNDVENOPTKFTPIGLNEVANGHFEQWTLPRY 408
 Db 358 RAHEAKVDTAGRTFKLSLEISTESSDFDQDQPTFRFPVGIG--VDNEADPQQWLSLPDY 416
 Qy 409 SGALTLMNMLAPAVAPLFGERLLPFRSVVPLKGGFGNPAIDCSVPQEWVQHFYQESAPS 468
 Db 417 SQQFTHMNMLAPAVAPNFGEQLLFRSOLPSSGGRSNGVLDCLVPQEWVQHFYQESAPA 476
 Qy 469 LGDVALVRYVNPDTGRVLFPEAKLHKGGLTWSSTSTGPVVPANGYFKFDSWVNOFYSLA 528
 Db 477 QTQVALVRYVNPDTGRVLFPEAKLHKGGLTWSSTSTGPVVPANGYFKFDSWVNOFYSLA 536
 Qy 529 PMGTGNGR 536
 Db 537 PMGTGNGR 544

RESULT 6
 AAB49708
 ID AAB49708 standard; protein; 542 AA.
 XX
 AC AAB49708;
 DT 04-APR-2001 (first entry)
 XX Small round structured virus protein SEQ ID 9.
 DE Small round structured virus; SRSV; food poisoning.
 KW Small round structured virus;
 KW Small round structured virus;
 OS Small round structured virus.
 XX WO200079280-A1.
 XX 28-DEC-2000.
 XX
 XX 22-JUN-2000; 2000WO-JP0040095.
 XX
 XX 22-JUN-1999; 99JP-00175928.
 XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 XX (DENK-) DENKA SEIKEN KK.
 XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 PI WPI; 2001-080848/09.
 DR N-PSDB; AAF29149.
 XX
 XX Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX
 PS Claim 1; Page 59-61; 84pp; Japanese.
 XX This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies

CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 542 AA;
SQ
Query Match 67.9%; Score 1933; DB 4; Length 542;
Best Local Similarity 67.0%; Pred. No. 7.4e-171;
Matches 366; Conservative 76; Mismatches 94; Indels 10; Gaps 6;
QY 1 MKMASNDATPSNDGAGLVPESSNEAMALEPVPVGLASLAAPVTGQTNIIIDPWRTNPFVQAP 60
DB 1 MKMASNDATPSNDGAGLVPESSNEAMALEPVPVGLASLAAPVTGQTNIIIDPWRTNPFVQAP 60
QY 61 NGEFTVSPRNSPGEILLVNLGPELNPYLALHARMYNGVAGMEVQVLMAGNAFTAGKII 120
DB 61 NGEFTVSPRNSPGEILLVNLGPELNPYLALHARMYNGVAGMEVQVLMAGNAFTAGKII 120
QY 121 FAAPVPYFVENLSPSQITMFPVHVIDVRLPVLPMDFVRLSTLPHFNQKDEPKMLVA 180
DB 121 FAAPVPYFVENLSPSQITMFPVHVIDVRLPVLPMDFVRLSTLPHFNQKDEPKMLVA 180
QY 181 MLYTLRNSGDDVFTVSCRILTRPSPEDFTYLVPPVTSKTKPFTLPVLTLGELSNS 240
DB 181 MLYTLRNSGDDVFTVSCRILTRPSPEDFTYLVPPVTSKTKPFTLPVLTLGELSNS 240
QY 241 RPLSIDEMVTSPNESIVVQONGRVTLDGELLGTTLOACNICISIRGVKVTQVPSQHM 300
DB 241 RPLSIDEMVTSPNESIVVQONGRVTLDGELLGTTLOACNICISIRGVKVTQVPSQHM 300
QY 301 WN---LEITNLNGTQDPTDDVPAPLGVDPDFAGEVGLSQNRNGES-NPANRAHDVA 474
DB 301 WN---LEITNLNGTQDPTDDVPAPLGVDPDFAGEVGLSQNRNGES-NPANRAHDVA 474
QY 475 VRYVNPDTGRVLFEAKLHKGGLTYSSTSTGVPVVPANGYFKFDSWVNOFYSLAPMGTCN 534
DB 475 VRYVNPDTGRVLFEAKLHKGGLTYSSTSTGVPVVPANGYFKFDSWVNOFYSLAPMGTCN 534
QY 535 GRRVQ 540
DB 537 GRRIQ 542
RESULT 7
ID AAB49704
XX AAB49704 standard; protein; 539 AA.
AC AAB49704;
XX
DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 5.
XX
XX Small round structured virus; SRSV; food poisoning.
XX
XX Small round structured virus.
OS
XX WO200079280-A1.
XX
XX 28-DEC-2000.
XX
XX 22-JUN-2000; 2000WO-JP004095.
XX
XX

PR 22-JUN-1999; 99JP-00175928.
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
XX
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
DR WPI; 2001-080848/09.
DR N-PSDB; AAF29145.
XX
PT Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
PS Claim 1; Page 50-52; 84pp; Japanese.
XX
CC This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX
SQ Sequence 539 AA;
Query Match 65.8%; Score 1872.5; DB 4; Length 539;
Best Local Similarity 64.5%; Pred. No. 3.2e-165;
Matches 349; Conservative 80; Mismatches 105; Indels 7; Gaps 4;
QY 1 MKMASNDATPSNDGAGLVPESSNEAMALEPVPVGLASLAAPVTGQTNIIIDPWRTNPFVQAP 60
DB 1 MKMASNDATPSNDGAGLVPESSNEAMALEPVPVGLASLAAPVTGQTNIIIDPWRTNPFVQAP 60
QY 61 NGEFTVSPRNSPGEILLVNLGPELNPYLALHARMYNGVAGMEVQVLMAGNAFTAGKII 120
DB 61 NGEFTVSPRNSPGEILLVNLGPELNPYLALHARMYNGVAGMEVQVLMAGNAFTAGKII 120
QY 121 FAAPVPYFVENLSPSQITMFPVHVIDVRLPVLPMDFVRLSTLPHFNQKDEPKMLVA 180
DB 121 FAAPVPYFVENLSPSQITMFPVHVIDVRLPVLPMDFVRLSTLPHFNQKDEPKMLVA 180
QY 181 MLYTLRNSGDDVFTVSCRILTRPSPEDFTYLVPPVTSKTKPFTLPVLTLGELSNS 240
DB 181 MLYTLRNSGDDVFTVSCRILTRPSPEDFTYLVPPVTSKTKPFTLPVLTLGELSNS 240
QY 241 RPLSIDEMVTSPNESIVVQONGRVTLDGELLGTTLOACNICISIRGVKVTQVPSQHM 300
DB 241 RPLSIDEMVTSPNESIVVQONGRVTLDGELLGTTLOACNICISIRGVKVTQVPSQHM 300
QY 301 WNLEITNLNGTQDPTDDVPAPLGVDPDFAGEVGLSQNRNGESNPANRAHDVA 474
DB 301 WNLEITNLNGTQDPTDDVPAPLGVDPDFAGEVGLSQNRNGESNPANRAHDVA 474
QY 475 VRYVNPDTGRVLFEAKLHKGGLTYSSTSTGVPVVPANGYFKFDSWVNOFYSLAPMGTCN 534
DB 475 VRYVNPDTGRVLFEAKLHKGGLTYSSTSTGVPVVPANGYFKFDSWVNOFYSLAPMGTCN 534
QY 535 GRRVQ 540
DB 537 GRRIQ 542
RESULT 8
ID AAB49710
XX AAB49710

DB	359	GGDHLCPQISSSEIYLTSPNLTCTNCPQLPQSLGRTILIRSDNGHCHDMVGTSPITPT	418
QY	380	ENQPTKFTPIGLNEVANGHRPEQWTLPRYSGALTINNNLAPAVAPLPGERLLFFRSVVP	439
DB	419	WPQWRRCRSGSGSCSGHRY-----PVPVWNRVTWIV-----LS	454
QY	440	LKGGFGNPAIDCSVPQ---EWQHFYQBSAPSLGDLVALRVYVNPDTGRVLFEAKLHKGF	496
DB	455	HKSGFSTST--RKLPLQNLRW-----PLIRFINPDTGRVLFEARLHKQGF	497
QY	497	LTVSSTSTGTPVVPVANGYFKFDSWVYQFYSYSLAPMTGNGRRVQ	540
DB	498	ITVAHTGDNPIVMPNGYFRFEAWVYQFYSYSLAPVGTGKGRRVQ	541
RESULT 9			
AA	50972	AAR50972 standard; protein; 530 AA.	
XX	AA	50972;	
DT	16-OCT-2003	(revised)	
DT	25-MAR-2003	(revised)	
DT	05-OCT-1994	(first entry)	
XX	Norwalk virus strain 8FIIa protein (encoded by ORF2).		
DE	XX	Norwalk virus; pathogen; acute gastroenteritis; food poisoning;	
KW	XX	seafood contamination; diagnostic assay; calcivirus; small round virus.	
XX	OS	Norwalk virus; (strain 8FIIa).	
XX	PN	WO9405700-A2.	
XX	PD	17-MAR-1994.	
XX	PF	07-SEP-1993; 93WO-US008447.	
XX	PR	07-SEP-1992; 92US-00941365.	
XX	PA	(BAYU) BAYLOR COLLEGE MEDICINE.	
XX	PI	Matson DO, Estes MK, Jiang X, Graham DY;	
XX	DR	WPI; 1994-101125/12.	
XX	DR	N-PSDB; AAQ56826.	
XX	PT	DNA from Norwalk and related viruses - used for preparing prods. for use in diagnostic assays, detection and vaccines for Norwalk and related viruses.	
XX	PS	Claim 14; Page 68-70; 156pp; English.	
XX	CC	The Norwalk virus was isolated from stool samples from adult volunteers infected with safety tested Norwalk virus strain 8FIIa. The coding sequence is useful for the design of probes for use in diagnostic assays for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 16-OCT-2003 to standardise OS field)	
XX	XX	Sequence 530 AA;	
Query Match 43.3%; Score 1233; DB 2; Length 530;			
Best Local Similarity 47.6%; Pred. No. 1.3e-105;			
Matches 263; Conservative 81; Mismatches 163; Indels 46; Gaps 11;			
QY	1	MMKASNDATPSNDGAAG---LVPESN-NEAMALEPVGASLAAPVGTQNIIDPWIRTNF	56
DB	1	MMWASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGVNPIDPILNIF	60
QY	57	VQAPNGEFTVSPRNSGCEILLVLELGPENLPYLAHLARMYNGYAGGMEVQVMLAGNAFTA	116
DB	61	VQAPQGEFTTSPNNTPGDVLFDLSLGLHLPNPFLLHLSQMYNGVGNVRVIMLAGNAFTA	120
Query Match 58.0%; Score 1650.5; DB 4; Length 541;			
Best Local Similarity 56.5%; Pred. No. 1.6e-144;			
Matches 330; Conservative 65; Mismatches 102; Indels 87; Gaps 8;			
QY	1	MMKASNDATPSNDGAAGLVPESNNEAMALEPVGASLAAPVGTQNIIDPWIRTNFVQAP	60
DB	1	MMWASNDAAPSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRTNFVQAP	60
QY	61	NCEFTVSPRNSGCEILLVLELGPENLPYLAHLARMYNGYAGGMEVQVMLAGNAFTAGKII	120
DB	61	AGEFTVSPRNSGCEILLVLELGPENLPYLAHLARMYNGYAGGMEVQVILVAGNAFTAGKII	120
QY	121	FAAVPPYPVPEVNSPSQITFMFHVHIDVTLEPVLPMDDYRSTLFHFNQDEPKMLVA	180
DB	121	FAAIPGPFYENLSPSQITMCPHVHIDVRQLEPFLPMDDIWNFFHYNQGNQDKPLRVA	180
QY	181	MLYTPLRNSGDDVFTVSCRILTRPSDFDTYLVPTVESKTPFTLPVLTGELSNS	240
DB	181	MLYTPLRANNSGDDVFTVSCRVLTKPSDFDTFLVPTVESKTFQALPILKISEMTNS	240
QY	241	RFPPLSIDEMVTSNIVVQPNQGRVTLGELLGTQLOACNICIRGKV---TQGVSE	297
DB	241	RFPVPVDMVTARNENQVQPNQGRVTLGELLGTTPLLAVNICFKPEVIKNGDVRVS-	299
QY	298	QHWNNLEITNLNGTQFDPTDDVPAPLGVDPDFAGEVGLVSQRNKGESNPANRAHDAVVAT	357
DB	300	-YRMDMEITNTGDTGDTPTEDTGPICGSPDFQGLFEGVASQRNKGESNPATRAHEAIINT	358
QY	358	YSDKYTPKGLGVQI-----GTWNTNDV	379

Wed Jun 2 09:13:32 2004

```

PD 28-DEC-2000.
XX
XX
PF 22-JUN-2000; 2000WO-JP004095.
PR
PR 22-JUN-1999; 99JP-00175928.
XX
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI: 2001-080848/09.
DR N-PSDB; AAF29142.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
XX Claim 1; Page 42-45; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
CC
XX Sequence 530 AA;
XX
XX Query Match 43.3%; Score 1232; DB 4; Length 530;
XX Best Local Similarity 47.3%; Pred. No. 1.6e-105;
XX Matches 264; Conservative 81; Mismatches 157; Indels 56; Gaps 12;
XX
QY 1 MKVASNDATPSNDGAAG---LVPESN-NEAMALEPVPVVGASLAAPVTGQTNIIDPWIRTNF 56
DB 1 MMWASKDAPTNDGTSVDSGAGQLVPEVNASDPLNDPVGASSTAVATAGQVNPIDPWIRNF 60
QY 57 VQAPNGEFTVSPRNSPGEILVNLELGPENLPYLAHLARMYNGVAGMEVQVMLAGNAFTA 116
DB 61 VQAPQGEFTISPNTPGGVLFDLGLPHLPNPLHLSQMYNGVGNMRVRIMLAGNAFTA 120
QY 117 GKIIIPAAPPVPFVENLSPSQITMFPVHVIDVRLTEPVLPMDDVRSITLPHFNQKDEPKM 176
DB 121 GKIIISCIPPGFAGHNLTIQAATLFPVHVIADVRLDFIEVPLEVDRVNLPHNDRNQQT 180
QY 177 RLVMALYTPLRSGDDVFTVSCRILTRPSPEFDFTYLVPPVTVESKTFPLPVLTLGE 236
DB 181 RLVCMLYTPLRGTGDSFVAGRVMTCPSPDFNFLVPPVTVEQKTRPFTLPNPLSS 240
QY 237 LNSRFPPLSIDEMVTSNPNESIVQPNQGRVTLDEGLLGTTOLOACNICSIRGKVTGOVPS 296
DB 241 LNSRAPLPISGMGISPNVQSVQPNQGRCTLDGLRLVGTTPVLSHVAKIRGTSNGTV-- 298
QY 297 EQHMNLEITNLNGTQDFDTPDDVPAPLGPVPDFAG-----EVEGVLSQRNRGESNPANR 349
DB 299 -----INLELDGTFPHFPFEG-PAPIGFPDLGGCDWHINTQFGHSSQTQ----- 342
QY 350 AHDVVATYSDKYTPKGLIVQ---IGTWNINDVENQPTKFTPIGLNEVA-----NGHRFE 401
DB 343 -YD--VDTPDPFVPHLGSIQANGIGSGNYIGV-----LSWVSPPSHSGSQVD 388
QY 402 QMTLPRYSALLTNMNLAPAVAPLPFGERLLFFRSVPLKGGFGNPAIDCSVPQWVQHF 461
DB 389 LMKIPNPGSSITEATHLAPSVYPPGGEVLVFPFMKIPGPGAYSIP---CLLPQEYISHL 445
QY 462 YQESAPSLGDVALRVVNPDTGRVLPEAKLHKGGLTV---SSTSTGPPVVVPANGYFKFDS 519
DB 446 ASEQATVGEAALLHVDPDTRTGLTGEFRAYPDGFLTCVPNGASSGPOQLPINGVVFVS 505
QY 520 WYNQFYSLAPMGTCNGRR 537
DB 506 WVSRYQLKPVGTASSAR 523

```

```

RESULT 12
AAB49700
ID AAB49700 standard; protein; 545 AA.
XX
AC AAB49700;
XX
XX 04-APR-2001 (first entry)
XX
XX Small round structured virus protein SEQ ID 1.
XX
XX Small round structured virus; SRV; food poisoning.
XX
XX Small round structured virus.
XX
XX WO200079280-A1.
XX
XX 28-DEC-2000.
XX
XX 22-JUN-2000; 2000WO-JP004095.
XX
XX 22-JUN-1999; 99JP-00175928.
XX
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI: 2001-080848/09.
XX N-PSDB; AAF29141.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX strains for investigation of food poisoning outbreaks, contains
XX antibodies.
XX
XX Claim 1; Page 40-42; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 545 AA;
XX
XX Query Match 43.1%; Score 1225; DB 4; Length 545;
XX Best Local Similarity 46.7%; Pred. No. 7.6e-105;
XX Matches 261; Conservative 89; Mismatches 173; Indels 36; Gaps 14;
XX
QY 1 MKVASNDATPSNDGAAG---LVPESN-NEAMALEPVPVVGASLAAPVTGQTNIIDPWIRTNF 56
DB 1 MMWASKDAPTNDGTSVDSGAGQLVPEANTAEPISEMEFVAGAAATAAGQVNNIDPWIRTNF 60
QY 57 VQAPNGEFTVSPRNSPGEILVNLELGPENLPYLAHLARMYNGVAGMEVQVMLAGNAFTA 116
DB 61 VQAPQGEFTISPNTPGGVLFDLGLPHLPNPLHLSQMYNGVGNMRVRIMLAGNAFTA 120
QY 117 GKIIIPAAPPVPFVENLSPSQITMFPVHVIDVRLTEPVLPMDDVRSITLPHFNQKDEPKM 176
DB 121 GKIIISCIPPGFAAQNISIAQATMFPVHVIADVRLDFIEVPLEVDRVNLPHNDRNQQT 179
QY 177 RLVMALYTPLRSG--SGDDVFTVSCRILTRPSPEFDFTYLVPPVTVESKTFPLPVLTL 234
DB 180 RLVCMLYTPLRASGSSGTDPPFVIAGRVLTCSPPFSFLVPPNVQKTKFSPVNLPL 239
QY 235 GELNSRFPPLSIDEMVTSNPNESIVQPNQGRVTLDEGLLGTTOLOACNICSIRGKVTGOV 294
DB 240 NTLNSRFPPLSIDEMVTSNPNESIVQPNQGRVTLDEGLLGTTOLOACNICSIRGKVTGOV 299
QY 295 PSEQHMNLEITNLNGTQDFDTPDDVPAPLGPVPDFAGEVGVLSQRNRGESNPANRAHDAV 354

```

Db 300 GNGY-----NLTELDGSPHAPES-PAPIGFPDL-GECDWHM-----EASPTTQFTNGD 347
 QY 355 V-----ATYSKYTKPLGLVQLGTWNTNDVENOPTKFTPIG-LNEVANGHR--FEQWTLP 406
 Db 348 VIKQINVKQESAFAPHLGTIQAD--GLSDSVNTNMIAXLGVSPVSDGHRGDVDPWIP 405
 QY 407 RYSGALTLMNMLAPAVLPPGERLLFFRSYVPLKGGFGNPAIDCSVPQEWVOHFYQBSA 466
 Db 406 RYGSTLVEAQAAPPYPPGFEAIVFFMSDFPIAHGTNGLSVPCCTIQOEFVTFVNEQA 465
 QY 467 PSIGDLVALRVYNPDGTGRVLFKAKLHKGGLTV--SSTSTGVPVVPANGYFKFDSWVQPF 524
 Db 466 PRGGAALLHYLDDPDRNLGFBKLYPEGFMTCVPSNCGTGPOTLPINGVFVFSWVSRF 525
 QY 525 YSLAPMGTCN-----GRRR 538
 Db 526 YQLKPVGTAGPACRLGIR 544

RESULT 13

AAR57091

ID AAR57091 standard; protein; 530 AA.

XX

AC AAR57091;

XX 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 05-OCT-1994 (first entry)

XX

DE Small round virus SRSV/KY/89 capsid protein.

XX

KW pathogen; acute gastroenteritis; food poisoning; seafood contamination;

KW diagnostic assay; human calcivirus; small round virus; SRSV; KY89;

KW Norwalk virus; capsid protein.

XX

OS Small round structured virus.

XX

PN WO9405700-A2.

XX 17-MAR-1994.

XX 07-SEP-1993; 93WO-US008447.

XX 07-SEP-1992; 92US-00941365.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX

XX Matson DO, Estes MK, Jiang X, Graham DY;

XX

XX WPI; 1994-101125/12.

XX N-PSDB; AAQ56832.

XX

XX DNA from Norwalk and related viruses - used for preparing prods. for use

XX

XX in diagnostic assays, detection and vaccines for Norwalk and related

XX

XX viruses.

XX

XX Example 7; Fig 13a; 156pp; English.

XX

CC The known sequence for Norwalk virus was used to obtain the sequence of
 CC other Norwalk-related viruses such as SRSV/KY/89, an agent from a stool
 CC from an outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide
 CC cDNA sequence includes part of the polymerase region and the capsid
 CC region of the genome; the deduced amino acid sequences are AAR57092 and
 CC AAR57091, respectively. Expression of fragments and derivs. of Norwalk-
 CC related viruses permits development of diagnostic assays to detect
 CC antibodies, antigens, viral genetic material or antivirals. (Updated on
 CC 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS
 CC field.)

XX Sequence 530 AA;
 SQ

Query Match 41.6%; Score 1183; DB 2; Length 530;

Best Local Similarity 46.3%; Pred. No. 6e-101;
 Matches 259; Conservative 84; Mismatches 158; Indels 58; Gaps 14;
 QY 1 MKMASNDATPSNDGAAG--LVPESN-NEAMALEPVPVGASLAAPVTGOTNIIDPWIRTNF 56
 Db 1 MMASKDATSSVDGASASVQLVPEVNASDPLAMDPAVGSSTAVATAGQVNPDPHIIINF 60
 QY 57 VOAPNGEFTVSPRNSPGSILNVLNLELGPBPNVLAHARMYNGYAGGMEVQVNLAGNAFTA 116
 Db 61 VOAPQGEFTISPNNTPGDVLFDLSLGPBPNVLAHARMYNGYAGGMEVQVNLAGNAFTA 120
 QY 117 GKIIIPAAVPVPVENLSPSQITWPHVYIDVRLTLEPVLMPDVRSTLPHNQKDEPKM 176
 Db 121 GKIIIVSCIPPGGSGOOLTIATLPPHVIADVRLTLEPVLMPDVRSTLPHNQKDEPKM 180
 QY 177 RLVAAMLYTLPSRNSGDDVFTVSCRILTRPSPEFDTYLVPTVPSKTKPFTLPVLTIGE 236
 Db 181 RLVCMLYTLSTGGTGDGDFVAVGRVMTCPSPDFNFLVLPPTVQKTRPFTLPNLPSS 240
 QY 237 LSNRPFLSIDEMVTSNPNESIVVQPNQGRVTLDSLLGTTQLOACNICSIRKVTGQVPS 296
 Db 241 LSNRAPLPISGMGISPDNVQSVQPNQGRVTLDSLLGTTQLOACNICSIRKVTGQVPS 298
 QY 297 EOHMNLITNLNGTQDPDTPDDVPAPLGVDPFAG-----EVFGVLQNRNGESNPANR 349
 Db 299 -----INLTLDGTPPHFEG-PAPIGFPDLGGCDWHINMTQFGHSSQTQ----- 342
 QY 350 AHDAVVATVSDKYTPKGLGVQ---IGTWNNDVENOPTKFTPIGLNEVA-----NGHRE 401
 Db 343 -YD--VDTPTDTSVPHLSIQANGIGSGNYIGV-----LSWSPSPHSGSQVD 388
 QY 402 QMTLPYSGALTLNMLNLAAPV-APLPPGERLLFFRSYVPLKGGFGNPAIDCSVPQSWVOH 460
 Db 389 LMKIENYSSITEATHLAPSVVSPGF-GEVLVFPMSKIP---GPGGDSLCLLPQGYISH 444
 QY 461 FYQESAPSLGDVALRVYNPDGTGRVLFKAKLHKGGLTV--SSTSTGVPVVPANGYFKFDP 518
 Db 445 LASEQAPTVEGEPGLLHYVDPDTRNLGFKAYPDGFLTCVPNGASSGPOQLPINGVFVFPV 504
 QY 519 SWVNQFYSIAPMGTCNGRR 537
 Db 505 SWVSRYQLKPVGTASTAR 523

RESULT 14

AAB49703

ID AAB49703 standard; protein; 544 AA.

XX

AC AAB49703;

XX

XX 04-APR-2001 (first entry)

XX

XX Small round structured virus protein SEQ ID 4.

XX

XX Small round structured virus; SRSV; food poisoning.

XX

XX Small round structured virus.

XX

XX WO200079280-A1.

XX

XX 28-DEC-2000.

XX

XX 22-JUN-2000; 2000WO-JP004095.

XX

XX 22-JUN-1999; 99JP-00175928.

XX

XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

XX

XX (DENK-) DENKA SEIKEN KK.

XX

XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

XX

XX WPI; 2001-080848/09.

XX

XX N-PSDB; AAF29144.

XX

```
XX PT Kit for the detection and typing of small round-structured virus (SRSV)
XX PN strains for investigation of food poisoning outbreaks, contains
XX PT antibodies.
XX
XX PS Claim 1; Page 47-49; 84pp; Japanese.
XX
XX CC This invention relates to a kit for the detection and typing of small
XX CC round structured virus (SRSV) strains. The kit contains antibodies
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX CC used for detecting and typing strains of SRSV in order to prevent the
XX CC spread of infection and to examine the epidemiology of outbreaks
XX
XX SQ Sequence 544 AA;
XX
XX Query Match 41.1%; Score 1170.5; DB 4; Length 544;
XX Best Local Similarity 45.6%; Pred. No. 9.2e-100; Indels 53; Gaps 16;
XX Matches 258; Conservative 85; Mismatches 170;
XX
XX QY 1 MKMASNDATPSNDGAAG--LVPESEN-EAMALEPVVGASLAAPVTGQNIIDPWIRTNF 56
XX DB 1 MMWASKDAPSDAGATGAGLVPEVNTADPLPEFVAGTAVATAGQVNLIDPWVNNP 60
XX
XX QY 57 VOAPNGEFTVSPRNSPGEILVNLELGPENLYLAHARMYNGVAGMEVQVMLAGNAFTA 116
XX DB 61 VOAPQGEFTISPNTPGDILFDLQGLPHNPFSLHSMYNGVGNMRVRLVLAGNAFTA 120
XX
XX QY 117 GKIIFAAVPPYFPVENLSPSQITMFPFHVIIIDVRLPEVLLPMDVRSITLHFNQKDEPKM 176
XX DB 121 GKVIICCVFPFQSRITLSIAQATLPHVIAADVRLDPVEPLEDVRLVLYH-NNDTQPTM 179
XX
XX QY 177 RLVMALYTPLRNSGS--GDDVFTVSCRILTRPSPEFDFTYLVPPTVESKTPFTLPVLTL 234
XX DB 180 RLICMLYTPLRNSGS--GDDVFTVSCRILTRPSPEFDFTYLVPPTVESKTPFTLPVLTL 239
XX
XX QY 235 GELNSRFPPLSIDMWTSPNESIVVQPNQGRVTLDELLGTTLQACNICIRGKVTQGV 294
XX DB 240 KYLSNRIPNPIEGMSLSPDQTVNQVFQNGRGTIDGQPLGTTFVSQKCKFRGIT--- 296
XX
XX QY 295 PSEOHMNLITNLNGTQFDPTDDVPAPLGVDPDFAGEVFGVLSQRNRGESNPANRAHDAV 354
XX DB 297 -SGORVNL--TELDGSPF-MAPAAPAGPFDLGSCDWHIEMSKIPNSSTQNNPIVTS 352
XX
XX QY 355 VATYSDKYTPKGLV-----QIGT--WNT--NDVENQTKFTPIGLNEVANGHR 399
XX DB 353 VKPNSQQFVPHLSITLDENSSGGDYIGTQWTSPPSDGGANTNF----- 399
XX
XX QY 400 FEOWTLPRYSGLATLNMNLAPAVAPLPFG--ERLLFFRSYVPLKGGFGNP-AIDCSVPQE 456
XX DB 400 ---WKIPDYGSLAASQALAPVY--PGFNEVYVFWASIPGPNQSGSPNLVPCLLPQE 454
XX
XX QY 457 WYQHFYQBSAPSLGDVALRVYVNDPTGRVLPFEAKLHKGGLTV--SSTSTGVPVVPANGY 514
XX DB 455 YITHFISEQAPIQGEAALLHYVDPDTRNRLGFEKLYPGGYLTCVPNSSTGPOQLPLDGV 514
XX
XX QY 515 FKFDWVWQFSLAPMGT-GNGRRV 539
XX DB 515 FVFASWVSRYQLKPVGTAGPARGRL 540
XX
XX RESULT 15
XX AAB49702
XX ID AAB49702 standard; protein; 546 AA.
XX
XX AC AAB49702;
XX
XX XX
XX DT 04-APR-2001 (first entry)
XX
XX DE Small round structured virus protein SEQ ID 3.
XX
XX KW Small round structured virus; SRSV; food poisoning.
```

```
XX OS Small round structured virus.
XX PN WO200079280-A1.
XX XX
XX PD 28-DEC-2000.
XX
XX PF 22-JUN-2000; 2000WO-JP004095.
XX
XX PR 22-JUN-1999; 99JP-00175928.
XX
XX PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX PA (DENK-) DENKA SEIKEN KK.
XX
XX PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX DR WPI; 2001-080948/09.
XX DR N-PSDB; AAF29143.
XX
XX PT Kit for the detection and typing of small round-structured virus (SRSV)
XX PT strains for investigation of food poisoning outbreaks, contains
XX PT antibodies.
XX
XX PS Claim 1; Page 45-47; 84pp; Japanese.
XX
XX CC This invention relates to a kit for the detection and typing of small
XX CC round structured virus (SRSV) strains. The kit contains antibodies
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX CC used for detecting and typing strains of SRSV in order to prevent the
XX CC spread of infection and to examine the epidemiology of outbreaks
XX
XX SQ Sequence 546 AA;
XX
XX Query Match 41.1%; Score 1169.5; DB 4; Length 546;
XX Best Local Similarity 44.9%; Pred. No. 1.1e-99;
XX Matches 254; Conservative 87; Mismatches 178; Indels 47; Gaps 13;
XX
XX QY 1 MKMASNDATPSNDGAAG--LVPESEN-EAMALEPVVGASLAAPVTGQNIIDPWIRTNF 56
XX DB 1 MMWASKDAPQADGASGAGQLVPEVNTADPLPEFVAGTAVATAGQVNLIDPWVNNP 60
XX
XX QY 57 VOAPNGEFTVSPRNSPGEILVNLELGPENLYLAHARMYNGVAGMEVQVMLAGNAFTA 116
XX DB 61 VOAPQGEFTISPNTPGDILFDLQGLPHNPFSLHSMYNGVGNMRVRLVLAGNAFTA 120
XX
XX QY 117 GKIIFAAVPPYFPVENLSPSQITMFPFHVIIIDVRLPEVLLPMDVRSITLHFNQKDEPKM 176
XX DB 121 GKIIICCVFPFGTSSSLTIAQATLPHVIAADVRLPEVPLEDVRLVLYHND-NQPTM 179
XX
XX QY 177 RLVMALYTPLRNS--NGSGD-DVFTVSCRILTRPSPEFDFTYLVPPTVESKTPFTLPVLTL 234
XX DB 180 RLVMALYTPLRNS--NGSGD-DVFTVSCRILTRPSPEFDFTYLVPPTVESKTPFTLPVLTL 239
XX
XX QY 235 GELNSRFPPLSIDMWTSPNESIVVQPNQGRVTLDELLGTTLQACNICIRGKVTQGV 294
XX DB 240 QTLNSRFPPLSIDMWTSPNESIVVQPNQGRVTLDELLGTTLQACNICIRGKVTQGV 295
XX
XX QY 295 PSEOHMNLITNLNGTQFDPTDDVPAPLGVDPDFAGEVFGVLSQRNRGESNPANRAHDAV 354
XX DB 296 --NQGARTLNLTEVDGKPFMAFDS-PAPVGFDPFGKCDWHMRISKTPNNTSSGDPMSVS 352
XX
XX QY 355 VATYSDKYTPKGLVQIGTWTNDVENVQPT-----KFTPIGLNEVANGHRPQE 402
XX DB 353 VQTNVQGFVPHLSIQF-----DEVFNHPTGDIYGTIEWISQPSPTPGTD-----INL 400
XX
XX QY 403 WTLPRYSGLATLNMNLAPAVAPLPFG--ERLLFFRSYVPLKGGFGNP-AIDCSVPQEWQHF 461
XX DB 401 WEIPDYGSSLSQAANLAPPVFPFGFGEALVYFSAFFGPNRNSAPNDVPCLLPQEVITF 460
XX
XX QY 462 YQBSAPSLGDVALRVYVNDPTGRVLPFEAKLHKGGLTV--SSTSTGVPVVPANGYPKFDS 519
```


Db 461 VSEQPTMGDAALLHYVDDTNRNLGEFKLYPGGYLTCVPNGVGAGPQOLPLNGVFLFVS 520

Qy 520 WYNQFYSLAPMGYGN-----GRRRV 539

Db 521 WVSRYQLKPVGTASTARSRLGVRR 546

Search completed: June 1, 2004, 13:45:56
Job time : 48.3893 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 12.8658 Seconds
(without alignments)
2166.837 Million cell updates/sec

Title: US-09-926-799-7
Perfect score: 2845
Sequence: 1 MKMASNDATPSNDGAAGLVP.....VNQFYSLAPMTGNGRRRVQ 540

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/prodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCTUS COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1233	43.3	530	4 US-08-486-049-3	Sequence 3, Appli
2	296.5	10.4	668	4 US-09-617-594A-4	Sequence 4, Appli
3	294.5	10.4	626	4 US-09-590-020-7	Sequence 7, Appli
4	281	9.9	622	4 US-09-590-020-6	Sequence 6, Appli
5	281	9.9	623	4 US-09-590-020-4	Sequence 4, Appli
6	277	9.7	623	4 US-09-590-020-2	Sequence 2, Appli
7	273	9.6	669	4 US-09-617-594A-2	Sequence 2, Appli
8	137	4.8	2318	3 US-09-091-219-24	Sequence 24, Appli
9	137	4.8	2318	4 US-09-660-541-24	Sequence 24, Appli
10	111.5	3.9	2206	1 US-07-852-260-2	Sequence 2, Appli
11	111.5	3.9	2206	2 US-08-461-503-2	Sequence 2, Appli
12	111.5	3.9	2206	3 US-08-465-250-2	Sequence 2, Appli
13	108.5	3.8	1125	4 US-09-513-783A-152	Sequence 152, App
14	108.5	3.8	1610	4 US-09-513-783A-22	Sequence 22, Appli
15	105.5	3.7	776	4 US-09-489-039A-13081	Sequence 13081, A
16	104.5	3.7	627	4 US-09-345-473B-46	Sequence 46, Appli
17	104.5	3.7	968	4 US-09-228-986-76	Sequence 76, Appli
18	104.5	3.7	2232	3 US-09-091-219-25	Sequence 25, Appli
19	104.5	3.7	2232	4 US-09-660-541-25	Sequence 25, Appli
20	104.5	3.7	2247	3 US-09-091-219-2	Sequence 2, Appli
21	104.5	3.7	2247	4 US-09-660-541-2	Sequence 2, Appli
22	103	3.6	519	4 US-09-489-039A-7785	Sequence 7785, Ap
23	101.5	3.6	878	1 US-08-237-919-2	Sequence 2, Appli
24	101.5	3.6	878	4 US-08-732-429-2	Sequence 2, Appli
25	101.5	3.6	878	4 US-09-798-267-2	Sequence 2, Appli
26	101.5	3.6	878	4 US-09-798-267-3	Sequence 3, Appli
27	101.5	3.6	878	5 PCT-US95-05518-2	Sequence 2, Appli

28 101.5 3.6 2736 4 US-09-252-991A-30227 Sequence 30227, A
29 100 3.5 853 4 US-09-489-039A-11009 Patent No. 11009, A
30 100 3.5 1091 6 5516630-2 Patent No. 5516630
31 100 3.5 2227 3 US-08-475-886-2 Sequence 2, Appli
32 100 3.5 2227 3 US-08-475-886-4 Sequence 4, Appli
33 100 3.5 2227 3 US-08-475-886-6 Sequence 6, Appli
34 100 3.5 2227 3 US-08-397-232-2 Sequence 2, Appli
35 100 3.5 2227 3 US-08-397-232-4 Sequence 4, Appli
36 100 3.5 2227 3 US-09-171-387-2 Sequence 2, Appli
37 100 3.5 2227 4 US-09-653-499-2 Sequence 4, Appli
38 100 3.5 2227 4 US-09-653-499-4 Sequence 6, Appli
39 100 3.5 2227 4 US-09-653-499-6 Sequence 2, Appli
40 100 3.5 2227 4 US-10-135-988-2 Sequence 2, Appli
41 100 3.5 2227 4 US-10-135-988-4 Sequence 4, Appli
42 100 3.5 2227 4 US-10-135-988-6 Sequence 6, Appli
43 99.5 3.5 1043 3 US-08-928-361B-30 Sequence 30, Appli
44 99.5 3.5 1721 3 US-08-700-651-5 Sequence 5, Appli
45 99.5 3.5 1721 3 US-08-928-361B-6 Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-08-486-049-3
; Sequence 3, Application US/08486049
; Patent No. 6572862
; GENERAL INFORMATION:
; APPLICANT: Estes, Mary K
; APPLICANT: Jiang, Xi
; APPLICANT: Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Ave., N.W.
; CITY: Washington, D.C.
; STATE:
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,049
; FILING DATE: June 7, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311.023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-486-049-3

Query Match 43.3%; Score 1233; DB 4; Length 530;
Best Local Similarity 47.6%; Pred. No. 2.5e-111;
Matches 263; Conservative 81; Mismatches 163; Indels 46; Gaps 11;
QY 1 MKMASNDATPSNDGAAG---LYPESN-NEAMALEPVVAGSLAAPTGTQNIIDPWIRTF 56
DB 1 MMASKDATTSSVDGASGAGQLVPEVNASDPLAMDVPVAGSTAVATAGQVNPIDPWIRTF 60

Qy	266	-----VTLDGELLGTTQLQACN-----ICSIRKVTQGV-----	294
Db	361	GVATDYIVPGIPDGWPDFTTIFGELIPAGDYAITNGTGDNDITATGYDTADIKNNTNFRG	420
Qy	295	-----PSEOHMNNLEITNLNGTQFDPDTPVAPLGVDPDFAGEVFGVLRSQRNRCESPANRA	350
Db	421	MYICGSLQRAWGDK--KISNTAFITT-----ATLDGNNNNKINPCNTI	461
Qy	351	HDVVVYSDKYTPKLGVLQIGTWNNDVENOPTKFTPIGLNEVANGHRPEQ-----WTLP	406
Db	462	DQSKIYVVFQDAHWGKA-----QTSDDTTLALLGYT--GIGEAIGSRDRDVRVRAISTLP	512
Qy	407	RYSCALTLMNMLNPAVAPLFPQERLLFFRSYVPLKGGFCGNPAIDCSVPQEWQHFVQESA	466
Db	513	E-TGAR-----CGNHPIFYNSIKL--GVYIRSIDV-----FNSQIL	546
Qy	467	PSLGDVALRVYNPDTPGRVLFEAKLHKGGFTVTSSTSTGTPVVVPVANGYPKFD	518
Db	547	HTSRQLSLNHYLLLPDPSFAVYRIIDNSNGSWFPDIGDSDGFSFVGVSFGKLE	598

RESULT 4

```

US-09-590-020-6
; Sequence 6, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnius, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-6

```

RESULT 6

US-09-590-020-2
; Sequence 2, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Villis, Alvira
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-1446
; CURRENT APPLICATION NUMBER: US/09/590.020
; CURRENT FILING DATE: 2000-06-08

; PRIOR APPLICATION NUMBER: 60/138,484
 ; PRIOR FILING DATE: 1999-06-10
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 623
 ; TYPE: PR1
 ; ORGANISM: Feline calicivirus
 ; US-09-590-020-2

Query Match
 Best Local Similarity 9.6%; Score 277; DB 4; Length 623;
 Matches 100; Conservative 53; Mismatches 142; Indels 108; Gaps 17;

Qy 30 EPVVGASLAPVTGOTNIIDPW-IRTFVQAPNGEFTVSPRNSPGEILLVNLGDELNPY 88
 Db 100 EPSAQMSAAADMATGKSDVSEWFAFFSEHTSNVW-----STSETQKILFKQSLGFLANPY 155
 Qy 89 LAHLARMYNGVAGGMEYQVMVNLGNAGNAFTAGRIIPAAVPPYFPVENLSPSQITWPHVVIDV 148
 Db 156 LSHLUAKUYVWAGSGVEVRFISGSGVGGKLAALVVP--GIEPVQSTSMLOYPHVLFDA 213
 Qy 149 RTEPVLPLMPDVRSTLHFHNQKDEPKRLVAMLYTPLRNSGDDVFTVSC--RLITRP 206
 Db 214 ROVEPVIFATPDLRSLNYHL-MSDITDTSIVIMVYNDL-INPYANDTNSGCCIVTVEIKP 271
 Qy 207 SPEDTFYLYPPPTVESKTKPTLPVLTLGELSNRPFLSIDEMVTSFNESIVVQPONGRV 266
 Db 272 GPDFKPHLLKPP-----GSMLT-----HGVS 292
 Qy 267 TLDGELLGTTLOQACNICIRGKVTQVPSEQHWNLEITNLNGTQDFPTDDVPALGVP 326
 Db 293 PSD-----LIPKSSSLW-----IGNRHS--DITDFIIRPF--- 321
 Qy 327 DFAGSVFGV-----LSQRNCGESNPANRAHDAAVVATYSDKYTPKLGIVQIGT----- 373
 Db 322 ----VFQANRHDFNQTAGMSTPRFR---PITITVSESNMSKLG-IGVATDIYVGP 372
 Qy 374 --WNINDVENOPTKPTPLGLNEVANGHRFEQWTLPRYSGALTL 414
 Db 373 DGWPPDTTI----PEQUTPAGIYSITASNGTIVITTAAGYDAAEFI 412

RESULT 7
 US-09-617-594A-2
 ; Sequence 2, Application US/09617594A
 ; Patent No. 6541458
 ; GENERAL INFORMATION:
 ; APPLICANT: Audonnet, et al.
 ; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
 ; FILE REFERENCE: 454313-3151.1
 ; CURRENT APPLICATION NUMBER: US/09/617,594A
 ; CURRENT FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: 60/193,332
 ; PRIOR FILING DATE: 2000-03-30
 ; PRIOR APPLICATION NUMBER: France 00 01761
 ; PRIOR FILING DATE: 2000-02-11
 ; PRIOR APPLICATION NUMBER: France 99 09421
 ; PRIOR FILING DATE: 1999-07-16
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 669
 ; TYPE: PR1
 ; ORGANISM: Feline calicivirus
 ; US-09-617-594A-2

Query Match 9.6%; Score 273; DB 4; Length 669;
 Best Local Similarity 22.4%; Pred. No. 1.3e-17;
 Matches 120; Conservative 74; Mismatches 188; Indels 154; Gaps 23;

Qy 67 SPRNSPGEILLVNLGDELNPYLAHLARMYNGVAGGMEYQVMVNLGNAGNAFTAGRIIPAAVPP 126

180 STSETQKILFKQSLGPLLNPYLEHLSKLYAVWSSGSDVDFRFSISGSGVFGGKLAALVVPV 239

127 YFPVENLSPSQITWFPVHIIIVRTLEPVLPMPPDVRSTLTPEHNOKQDPKMRVLVAMLYTPL 186

240 --GVDPVOSTMLOVPHVLFQARQVPIPSIDPLRSTLYHL-MSDDTTTSLVIMVNDL 296

187 -----RNGSGDDVFTVSCRILTRPSPFEFTVLYVPTVESKTKPPTLVLTGLGELS 240

297 INPYANDSNSG-CIVTVE---TKPGPDFKFKLLKPPG-----SMLTHGSI PSD 341

241 RPPLS-----IDEMVTSNHSIWQP-----ONCR-----VTL---DG 270

342 LIPKSSSLWIGNRYWSDITDFVIRPFVQANRHDFNQETAGNSTPFRPTITISESNG 401

271 ELLGTTQLQACNICSIIRGKAVTQGVPSQHW-----NLEITNLNGTQ----- 312

402 SKLGT-----GVATDIYVPGIDGWPDPDTTIGEELTPAGDYSITNGSGNDIATA 449

313 --FDPTDDVPAPLGVDPDFAGEVFGVLSQRNGES-----NPNRAHD 352

450 NAYDSADVI---TNTTFRMGYICGALQARMGDKKISSTAFITTAKEGNTLKESNIIDM 506

353 AVVATVSDKYTKPLGLVQIGTWTNDVENQPTKFTPI---GLNEVANGHREQWTLPRYS 409

507 TKIAYQDTHVGR-----DVQTSDDTLAILGYTGIGEOAIGNRDS----- 547

410 GALTLMNLAAPAVAPLPGERLLFFRSYVPLKGGFGNPAIDCSVPQEWVQHVFQESAPSL 469

548 ---VVRISMLPETGAR-GGNHPIFYKNSIKL--GYVLRSIDV-----FNSQILHTS 592

470 GDVALVRYNVP-----DTGRVLPKALKHKGGLTSTSTSGPVVVPANGYF 515

593 RQLSLNHYLLPPDSFAVYRIIDNSNGSWFDGIDSDGFSFVGSSPKLEFPPLSASY 648

RESULT 8

US-09-091-219-24

; Sequence 24, Application US/09091219

; Patent No. 6171592

; GENERAL INFORMATION:

; APPLICANT: STUDDERT, Michael J.

; APPLICANT: CRABB, Brendan S.

; APPLICANT: FENG, Li

; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS

; FILE REFERENCE: 040268/0151

; CURRENT APPLICATION NUMBER: US/09/091,219

; CURRENT FILING DATE: 1998-10-05

; EARLIER APPLICATION NUMBER: PCT/AU96/00815

; EARLIER FILING DATE: 1996-12-18

; EARLIER APPLICATION NUMBER: AU PN7201

; EARLIER FILING DATE: 1995-12-18

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 24

; LENGTH: 2318

; TYPE: PRT

; ORGANISM: Foot-and-mouth disease virus

US-09-091-219-24

Query Match 4.8%; Score 137; DB 3; Length 2318;

Best Local Similarity 21.6%; Pred. No. 0.0017;

Matches 77; Conservative 40; Mismatches 130; Indels 110; Gaps 14;

QY 11 SNDGAGLVLP-----ESNNE-----AWALEPVPVVGASLAAPVTGQTNIIDPWFRTNFVQA 59

```

QY      60  ENGEIVSRKNSFG-----LAKK-----
      ||| : : : ||| : : :
Db      304  RNGTHITSTTQSSGVTVGYATADEFVSGNTSGLETRVVQAEKRFKTHLPWVTSDFGR 363
      ||| : : : ||| : : :
QY      79  ---LELGPELNPYLAHARMNYGAGMEVQVMLAGNAFTAGKIIFAAVPVPFVENVLSLP 135
      ||| : : : ||| : : : ||| : : : ||| : : :

```

Db 364 CHLLELPDTHKGVYGLTSDSYAMRNGWDVEVTAVGNQFNQGLLVAMVPPELYSTQKREL 423
QY 136 SQITMPPHVIDVRLTLEPVLPMPPDVRSTLFHFNQKDEPKMR-----LVAMLYTPPLRSNGS 191
Db 424 YQLTLFPHQFINPRTNMTAHTVFPV-----GVNRYDQVKVHKPWTLVVMVVAPLTVNTE 478
QY 192 GDDVFTVSCRILTRSPEDFTYLVPPPTVESKTKPFTLVTLGELSNR--FPLSIDB- 248
Db 479 G-----APQIKVYANIAPT-----NVHAGEFPFSKEGIFPVACSDG 514
QY 249 ----MVTSPNESIVVQPNQGRV-----TLDCGELLGTTQLQACNICSIIRKVTGOVP 295
Db 515 YGGLVTTDPK---TADPVYGVKVFNPRLNQLGRF--TNLLDVAEACPTFLRPEGGVP 566

RESULT 9

US-09-660-541-24
; Sequence 24, Application US/09660541
; Patent No. 6531136
; GENERAL INFORMATION:
; APPLICANT: STUDDERT, Michael J.
; APPLICANT: CRABB, Brendan S.
; APPLICANT: FENG, Li
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
; FILE REFERENCE: 040268/0151
; CURRENT APPLICATION NUMBER: US/09/660,541
; CURRENT FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/091,219
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: AU PN7201
; PRIOR FILING DATE: EARLIER FILING DATE: 1995-12-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 2318
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
US-09-660-541-24

Query Match 4.8%; Score 137; DB 4; Length 2318;
Best Local Similarity 21.6%; Pred. No. 0.0017;
Matches 77; Conservative 40; Mismatches 130; Indels 110; Gaps 14;
QY 11 SNDGAAGLVP-----ESNNE-----AMALEPVVGASLAAPVTGQTNIIDPMTNFWQA 59
Db 248 SNEGSTDTTSTHTTQNDWFSKLASSAFSLGALLADKKTEETLLEDRLIT-----T 303
QY 60 NGEFTVSPRSPG-----EILVN----- 78
Db 304 RNHTTSTTQSSVGTGYATAEDFVSGPNTSGLETRVVQAEPRFKTLFDWVTSDFGR 363
QY 79 ---LELGPENLPYLAHARMYNGYAGGMEVQVMLAGNAFTAGKIIFAAVPPYFPVENLSP 135
Db 364 CHLLELPDTHKGVYGLTSDSYAMRNGWDVEVTAVGNQFNQGLLVAMVPPELYSTQKREL 423
QY 136 SQITMPPHVIDVRLTLEPVLPMPPDVRSTLFHFNQKDEPKMR-----LVAMLYTPPLRSNGS 191
Db 424 YQLTLFPHQFINPRTNMTAHTVFPV-----GVNRYDQVKVHKPWTLVVMVVAPLTVNTE 478
QY 192 GDDVFTVSCRILTRSPEDFTYLVPPPTVESKTKPFTLVTLGELSNR--FPLSIDB- 248
Db 479 G-----APQIKVYANIAPT-----NVHAGEFPFSKEGIFPVACSDG 514
QY 249 ----MVTSPNESIVVQPNQGRV-----TLDCGELLGTTQLQACNICSIIRKVTGOVP 295
Db 515 YGGLVTTDPK---TADPVYGVKVFNPRLNQLGRF--TNLLDVAEACPTFLRPEGGVP 566

RESULT 10

US-07-852-260-2
; Sequence 2, Application US/07852260
; Patent No. 5525715
; GENERAL INFORMATION:

; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/852,260
; FILING DATE: 19920619
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-B-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-852-260-2

Query Match 3.9%; Score 111.5; DB 1; Length 2206;
Best Local Similarity 17.9%; Pred. No. 0.47;
Matches 107; Conservative 74; Mismatches 170; Indels 247; Gaps 28;
QY 11 SNDGAAGLVP-----ESNNEAMALSPV-----VGASLAAPVT----- 42
Db 270 TNSGATVLPVYNALAIIDSMVKHNNWGIAILPLDPAQDSSVEIPITVTIAPMCSEFN 329
QY 43 GQTNLIIDP-----WIRTNFVQAPNG--EFTVSPR-NSPGEI----- 75
Db 330 GLRNVTAQFQGLPVLNTPGNSQYLTSNQHOSPCAIPFDVTPPIDIPGEVKNMMLAEI 389
QY 76 -----LVNLELGPENLPYLAH-----LARMYNG 98
Db 390 DTMIPLNLESTKRNMTDMYRVTLSDSADLSQILCLSLSPADPRLSHMTMLGEVLNYTH 449
QY 99 YAGGMEVQVMLAGNAFTAGKIIFAAVPPYFPVENLSPSQITMPPHVIDVRLTLEPVLPM 158
Db 450 WAGSLKFTFLFCGSMATKILVAYAPGAQPP--SRKEAMLGTHVINDLGLQSSCTMVV 508
QY 159 PDVSTSLPHFNQKDE-----PKMRLVAMLYTPPLRSNGSGDDVFTVSCRILTRSP 208
Db 509 PWISNVTVRQTQDSFTSGGYISMFYQTRIVVPLSTPKSMMLG---FVSACN----- 558
QY 209 EFDFTYLVPPPTVESKTKPFTLVTLGELSNRPPPLSIDEMVTSPNESIVVQPNQGRVTL 268
Db 559 DFSVRLLEDTT-----HISQALPQGIEDLT-----SEVAQ----- 589
QY 269 DCELLGTTQLQACNICSIIRKVTGOVPSEQHMMNLITNLNGTQDPDTPDDVPA----- 321
Db 590 -----GALTLSLPKQD--SLPDTKASGPAH--SKEVPALTAVETG 626
QY 322 ---PLGVDPDFAGEVGVLSQRNGES-----NPNRAHDVAV--- 355
Db 627 ATNPL-APSDTVQTRHVHVQRSSRSESTIESPFARGACVAAIEVDNEQPTTQAQKLFAWVR 685

QY	356	ATYSKDYTPKGLGVOIGTWTNDVENOPTKFTPI---	GLNEVANGHRFEQWTLPRYS GAL 412
DB	686	ITYKDTVOLRRKL-EFFTYSRFDM-	FTFVVNTANNGHALNQ-----729
QY	413	TLMNMLAPAVAPLPGERLLFFRSYVPLKGGFGNPAIDCSVPQEWQHIFYQESA-PSL 469	
DB	730	----	VYQIMYIPP-----GAP-----TPKSDDYTWTOTSSNP SI 759
RESULT 11			
US-08-461-503-2			
; Sequence 2, Application US/08461503			
; Patent No. 5834302			
GENERAL INFORMATION:			
APPLICANT: Racaniello, Vincent			
APPLICANT: Tatem, Joanne M.			
TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES			
TITLE OF INVENTION: FROM CDNA			
NUMBER OF SEQUENCES: 9			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Cooper & Dunham			
STREET: 1185 Avenue of the Americas			
CITY: New York			
STATE: New York			
COUNTRY: U.S.A.			
ZIP: 10112			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: Patent In Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/461.503			
FILING DATE: 5-JUN-1995			
CLASSIFICATION: 435			
ATTORNEY/AGENT INFORMATION:			
NAME: White, John P.			
REGISTRATION NUMBER: 28,678			
REFERENCE/DOCKET NUMBER: 36607-D-PCT-US			
TELEPHONE: (212) 278-0400			
TELEFAX: (212) 391-0525			
TELEX: 422523 COOP UI			
INFORMATION FOR SEQ ID NO: 2:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 2206 amino acids			
TYPE: amino acid			
TOPOLOGY: linear			
MOLECULE TYPE: protein			
US-08-461-503-2			
Query Match 3.9%; Score 111.5; DB 2; Length 2206;			
Best Local Similarity 17.9%; Pred. No. 0.47;			
Matches 107; Conservative 74; Mismatches 170; Indels 247; Gaps 28;			
QY	11	SNDGAAGLV-----ESNNEMALEPV-----VGASLAAPVT-----42	
DB	270	TNSATIVPYNALADSMVKHNGIAILPLSLDFAQDSSEIVPTITVIAWCSEFN 329	
QY	43	GQTNIIDP-----WRTNFQAPNG--EFTVSFR-NSPGEI-----75	
DB	330	GLRNVTA PKFGLPVLNTPGSNYLSHDHQSCAIPEDVTPPIDIGEVKNMMELAEI 389	
QY	76	-----LVNLELPENLYAH-----LARMYNG 98	
DB	390	DTMIPLNLESTKRNTMDMYRVLTSDSADLSQPILCLSLSPAFDPRLSHTMLGEVNLNYTH 449	
QY	99	YAGMEVQVMLAGNAFTAGTIIFAAVPPYPVENLSFSQITWPFPHVIDVRTLEPVLPM 158	
DB	450	WAGSLKFTFLFCGSGMMATGKILVAIYAPPQAQPT-SRKEAMLGTHVIWDLGQSCTWV 508	

Db 270 TNSATIVPYNALAIDSMVKHNNWGIALLPLSPLDFAQDSVEIPITVTIAPMCSEFN 329
QY 43 GQNIIDP-----WIRTFVQAPNG--EFTVSPP--NSPGEI-----75
Db 330 GLRNVTAPOGLPVLNTPGSNQYLTSDNHQSPCAIPEDFTVPIIDPGEVKNMELAEI 389
QY 76 -----LVNLELQPELNPYLAH-----LARMYNG 98
Db 390 DTMIPLNLESTKRNMTDMRYRTLSDSADISQPICLSLSPAFDPRLSHTMLGEVLNYTH 449
QY 99 YAGGMEVQVNLGNAFTAGKIIFAAVPPYFPVENLSPSQITWFFPHVILIDVRLTLEPVLPM 158
Db 450 WAGSLKFTFLFCGSMWATKILVAYAPGAQPP--SRKEAMLGTHVIMDLGQSSCTMVV 508
QY 159 PDVRSTLFHFNQDE-----PKMLVAMLYTPLRSNGSGDDVFTVSCRILTRPSP 208
Db 509 PWISNVYRQTQDSFTEGGYISMFYQTRIVVPLSTPKSMMLG---FVSACN-----558
QY 209 EFDFTVLVPTVESKTKPFTPLVLTGELSNRFPPLSIDEMVTSPNESIVQPOQNGRVTL 268
Db 559 DFSVRLLRDT-----HISQALPOGIEDLT-----SEVAQ-----589
QY 269 DCELLGTTLOACNICSIRGKVTGQVPSQHMWNLEITNLNGTQDPTDDVA-----321
Db 590 -----GALTSLUPKQOD--SLPDTKASGAH--SKEVPALTAVETG 626
QY 322 ---PLGVPDPAGEVGLVLSQRNGES-----NPNRAHDAVV---355
Db 627 ATNPL-APSDTVQTRHVVRSESTIESFFARGACVAIIEDVNEQPTTRAKLPAMWR 685
QY 356 ATYSDKYTPKGLVQIGTWTNDVNOPTKFTPI---GLNEVANGHRFEQWTLPRYSGAL 412
Db 686 ITYKDTVOLRRKL-EPFTYSRFDME-----FTFVVTANFTNANGHALNQ-----729
QY 413 TLNMLAPAVAPLFPGERLLFRSYVPLKGGCNPAIDCSVQEWVQHFYQBSA-PSL 469
Db 730 -----VYQIMYIPP-----GAP-----TPKSWDDYTWTQSSNPISI 759

RESULT 13
US-09-513-783A-152
; Sequence 152, Application US/09513783A
; Patent No. 6416959
; GENERAL INFORMATION:
; APPLICANT: Giuliano, Kenneth A.
; APPLICANT: Kapur, Ravi
; TITLE OF INVENTION: A System for Cell Based Screening
; FILE REFERENCE: 97-022-L1
; CURRENT APPLICATION NUMBER: US/09/513,783A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 152
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-513-783A-152

Query Match 3.8%; Score 108.5; DB 4; Length 1125;
Best Local Similarity 21.2%; Pred. No. 0.32;
Matches 130; Conservative 65; Mismatches 232; Indels 187; Gaps 29;
QY 5 SNDATPSNDGAGLVPSNNEAMALEPVVVGASLAAP-----VTG-----QT 45
Db 256 STDTTPGPDTEAALAKDI--BEITKPDVILANVTQPTSTESDMFLAQDMELLTGTEAAHAN 313
QY 46 NIIDPW-----IRTNFVQAP-----NGEFTVSPRNSPGIILVNLELGP-----EL 85
Db 314 NIILPTEPDESSTKDVAPPMEEIIVPGNDTTSFKETETTLPIKMDLAPPEDVLLTKETEL 373
QY 86 NPYLAHLARMYNGYAGCM-----EVQVMLAGNAFTAGKIIFAAVPPYFPVENLSPSQITWF 141

Db 374 AP-----AKGMVLSLSEIEALAKNDVRSABE1-----PVAQETVVSETEVLA 415
QY 142 PHVII---DVRTL-EPVLLPMPDVRSTLTFHFNQKDEPKMLRVAMLYTPLRSN--GSGDDVF 196
Db 416 TEVVLPSDPITTLTKDVTLPLEAERPLVTDMPSTLETEMTLGKETAPPTETNLGMAKDM- 474
QY 197 TVSCILIRPSPPEPDT-----YLVPPPTVESKTKPFTLPVLTGELSNRFPPLSIDEMVT 251
Db 475 -----SPLPESEVTILGKDVVILP---ETKVABF-----NNVTPLS-EEBVT 511
QY 252 SPNESIVVQPOQNGRVTLDGELGTLTQLOACN-----ICSIKGVKTVGOVPS 296
Db 512 SYKDNPSPAETAPLAKNADLHSGTELVINDSNMAPASDLALPLETKVATVPKDKGTVOT 571
QY 297 EQHMNLEITNLNGTQDPTDDVPAPLGVDP--PAGEVFGV-----LSQRNRGE 343
Db 572 BEK--PREDSQLASQHKQSTVPPCTASPEPVKAAEQMSTLPIDAPSPLENLEQKETPG 629
QY 344 SNPA-----NRAHDAVATYSDKYTP-----KGLVQIGTWTNDVE 380
Db 630 SQSPSPSGVSRQEBEAKAAVGTGNDITTPPNKPEPPSPKCAKPLATTQPAKTSTSKAK 689
QY 381 NQPTKFTPIGLNEVANGHRFEQWTLPRYSGALTLN-MNLAPAVAPLFPGER-----LLFF 434
Db 690 TQPTS-----LPKQAPATTSGGLNKKPMSLASGSVPAAPHKRPAATAATAR 735
QY 435 RSYVPLKGGFGNPAIDCSVQEWVQHFYQESAPSLGDAVALVRYVNPDTGRVLFPEAKLHKG 494
Db 736 PSTLPARDVKPKPITEAKVAEKRTSPSKPSSAPALKP-----GPKTTPTVSKA-----783
QY 495 GFLTVSST--STGP 506
Db 784 ---TSPSTLVSTGP 794

RESULT 14
US-09-513-783A-22
; Sequence 22, Application US/09513783A
; Patent No. 6416959
; GENERAL INFORMATION:
; APPLICANT: Giuliano, Kenneth A.
; APPLICANT: Kapur, Ravi
; TITLE OF INVENTION: A System for Cell Based Screening
; FILE REFERENCE: 97-022-L1
; CURRENT APPLICATION NUMBER: US/09/513,783A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: EYFPP-DEVD-MAP4-EBFP construct
US-09-513-783A-22

Query Match 3.8%; Score 108.5; DB 4; Length 1610;
Best Local Similarity 21.2%; Pred. No. 0.56;
Matches 130; Conservative 65; Mismatches 232; Indels 187; Gaps 29;
QY 5 SNDATPSNDGAGLVPSNNEAMALEPVVVGASLAAP-----VTG-----QT 45
Db 502 STDTTPGPDTEAALAKDI--BEITKPDVILANVTQPTSTESDMFLAQDMELLTGTEAAHAN 559
QY 46 NIIDPW-----IRTNFVQAP-----NGEFTVSPRNSPGIILVNLELGP-----EL 85
Db 560 NIILPTEPDESSTKDVAPPMEEIIVPGNDTTSFKETETTLPIKMDLAPPEDVLLTKETEL 619
QY 86 NPYLAHLARMYNGYAGCM-----EVQVMLAGNAFTAGKIIFAAVPPYFPVENLSPSQITWF 141
Db 620 AP-----AKGMVLSLSEIEALAKNDVRSABE1-----PVAQETVVSETEVLA 661

RESULT 15
 US-09-489-039A-13081
 ; Sequence 13081, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 13081
 ; LENGTH: 776
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-13081

	Query Match	3.7%	Score 105.5	DB 4	Length 776
Best Local Similarity	21.2%	Pred. NO. 0.34			
Matches 120; Conservative	63	Matches 218	Indels 165	Gaps 26	
19	VPESNNEAMALEPV-----VGASLAAPVPTGQTNIDPWRTNFVQAPNGEETVSPRN	70			
24	VESSEGIIVAKVPWMDKNPSRIGQSLPDAVTSQTRIRRPVACYLQ--HGPVSREGRG	81			
71	SPGSEILVNLELGPSELNLYLAHLAR-----WYNGYAG-----GMEYQV	107			
82	KEPFEVYSWEALDL-----LARELSVKARCNEAIYGGSYGWASAGRFFHAQSOLHR	135			
108	MLAG-NAFTAGKIIFAAVPPVFPVENLSQIITMFPHVIIDRTTLEPVLPMDPVRSITLF	166			
136	FLKGGGGTASTNYIS-----SAGERILPH-----ILGPLSLPHRQHT	174			
167	HFNQKDEPKMRLVAMLYTPURS-----NCGG-----DDVFTVTSCLITRPSPEEFT	213			
175	HFSELARCOLFVAIGLPLFNACVNGGGANDHMLQWLDKMQANTREI NISPRNDL	233			

214	QY	YLVPPTVESKTKFTPLVLTLGELNSRPLSIDEMVTPSNESI	VVQPQNGRVTLDBELL	273
234	Db	SAVFPDAEWLAIRCGTTALLA-LSVILTAESLYDQAFVASH	TVGFAPY--RAYLLGSHD	290
274	QY	GT-----TQLQACNIC-----SIRGKVTCQVPDSEOHM	-NLEITNLN	309
291	Db	GVAKTPAWAAITGLDQAQRIADLAREMARHRTWVAISIO	RARQGEQAYWATVALTALL	350
310	QY	QTQFPDPTDDVPALGVDPDFAGEFVGLSQNRGESNPA--	NRAHDAV--VATYSDKYTPKL	366
351	Db	GQLGTPGGGLGFGYACTNLGAYKAFS---GPRLPAGENA	VDSVIPVARLSDM----	401
367	QY	GLVQIGTWNTDNDVENQPTKPTGLNEVANGHRF-----	BQWTLPR-----Y	408
402	Db	--LLHPG--ETVEFGQORRPDIHLVYWAGENAFHHQDIN	RLCEAWRRPETVWVHEQYW	458
409	QY	SGALTLNNLAPAVAPLPGER-----LLFRSRYVPLKG-		442
459	Db	TAQAKFSDIVLPATTSL--EREDIGSGGHGDFWIAMSAQ	IPPVGEARDDYAFCDLAGR	515
443	QY	-GFGNPAIDCSVPQEWQHFQESAP	467	
516	Db	IGFGFEAFSEGRDAGOMLRLHLYEESRP	541	

Search completed: June 1, 2004, 13:58:02
Job time : 14.8658 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 34.5201 Seconds
(without alignments)
4368.312 Million cell updates/sec

Title: US-09-926-799-7
Perfect score: 2845
Sequence: 1 MKMASNDTPSNDGAGLVP.....VNQFYSLAPMTGNGRRRVQ 540

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1233	43.3	530	14	US-10-314-739-3
2	296.5	10.4	668	14	US-10-209-507-4
3	292.5	10.3	547	12	US-10-670-695-4
4	292.5	10.3	671	12	US-10-670-695-2
5	273	9.6	669	14	US-10-209-507-2
6	138	4.9	1147	15	US-10-327-481A-38
7	131.5	4.6	538	10	US-09-867-932A-2
8	117.5	4.1	5376	15	US-10-028-248A-74
9	117.5	4.1	5376	15	US-10-147-026-74
10	111	3.9	6304	14	US-10-107-026-16
11	109.5	3.8	689	14	US-10-207-706-2
12	109	3.8	257	12	US-10-282-122A-76092
13	108.5	3.8	1125	14	US-10-100-957A-152
14	108.5	3.8	1521	14	US-10-238-075-909
15	108.5	3.8	1539	14	US-10-085-959-253

16	108.5	3.8	1610	14	US-10-100-957A-22	Sequence 22, Appl
17	108.5	3.8	5374	15	US-10-028-248A-75	Sequence 75, Appl
18	108.5	3.8	5374	15	US-10-107-782-75	Sequence 75, Appl
19	107.5	3.8	1152	15	US-10-369-493-12027	Sequence 12027, A
20	107	3.8	1544	15	US-10-369-493-6078	Sequence 6078, Ap
21	106.5	3.7	882	14	US-10-097-340-37	Sequence 37, Appl
22	106.5	3.7	894	12	US-10-425-114-39106	Sequence 39106, A
23	104.5	3.7	561	12	US-10-282-122A-52830	Sequence 52830, A
24	104.5	3.7	627	9	US-09-862-027-46	Sequence 46, Appl
25	104.5	3.7	968	14	US-10-101-464A-76	Sequence 76, Appl
26	103.5	3.6	785	15	US-10-369-493-11883	Sequence 11883, A
27	103.5	3.6	3930	12	US-10-282-122A-46817	Sequence 46817, A
28	102.5	3.6	1501	12	US-10-282-122A-44608	Sequence 44608, A
29	101.5	3.6	878	9	US-09-905-983-48	Sequence 48, Appl
30	101.5	3.6	878	13	US-10-165-049-2	Sequence 2, Appl
31	101.5	3.6	878	13	US-10-165-049-3	Sequence 3, Appl
32	101.5	3.6	878	15	US-10-173-551-34	Sequence 34, Appl
33	101.5	3.6	2468	12	US-10-282-122A-66335	Sequence 66335, A
34	101.5	3.6	2468	14	US-10-246-330-4	Sequence 4, Appl
35	101	3.6	609	15	US-10-369-493-22725	Sequence 22725, A
36	101	3.6	609	15	US-10-369-493-22806	Sequence 22806, A
37	101	3.6	925	15	US-10-108-260A-3889	Sequence 3889, Ap
38	101	3.6	6310	12	US-10-282-122A-67793	Sequence 67793, A
39	100.5	3.5	903	15	US-10-369-493-14300	Sequence 14300, A
40	100.5	3.5	1147	15	US-10-369-493-11450	Sequence 11450, A
41	100.5	3.5	1151	15	US-10-369-493-14817	Sequence 14817, A
42	100.5	3.5	1151	15	US-10-369-493-15006	Sequence 15006, A
43	100.5	3.5	2358	12	US-10-282-122A-45763	Sequence 45763, A
44	100	3.5	578	9	US-09-924-358-39	Sequence 39, Appl
45	100	3.5	578	15	US-10-410-764-39	Sequence 39, Appl

ALIGNMENTS

RESULT 1

US-10-314-739-3

; Sequence 3, Application US/10314739

; Publication No. US20030129588A1

; GENERAL INFORMATION:

; APPLICANT: Estes, Mary K

; Jiang, Xi

; Graham, David Y

; TITLE OF INVENTION: Methods and Reagents to Detect and

; Characterize No. US20030129588A1walk and Related Viruses

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pubbright & Jaworski L.L.P.

; STREET: 801 Pennsylvania Ave., N.W.

; CITY: Washington, D.C.

; STATE: <Unknown>

; COUNTRY: USA

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/314,739

; FILING DATE: 09-Dec-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,049

; FILING DATE: June 7, 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Davis, Peter

; REGISTRATION NUMBER: 36,119

; REFERENCE/DOCKET NUMBER: 311.023

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-662-0200

; TELEFAX: 202-662-4643

; TELEX: <Unknown>

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-4

Query Match          10.4%; Score 296.5; DB 14; Length 668;
Best Local Similarity 26.0%; Pred. No. 8.1e-19;
Matches 114; Conservative 55; Mismatches 135; Indels 135; Gaps 22;

QY 11 SMDGAGLVPESSNEAMALEPVGASLAAPVTGQNTIIPWIRTNFVQAPNGEFTV--- 66
Db 125 ADDGSV-TTPEQQT-----LVGGVIAEPNAQMSAVAD--VATG--KSVSEWEAFPSF 172
QY 67 -----SPRNSPGEILVNLELPELNPYLALHARMYNGVAGMEVQVMLAGNAFTAGKII 120
Db 173 HTSVNWTSETQKILFKQSLGPLNLPYLTHLAKLYVAWSGSIEVRFSISGSGVFGKLA 232
QY 121 FAAVPPVF-PVENLSPSQITMFPFHVIIIVRTLEPVLPMVDVRSSTLPHFNQKDEPKMRLV 179
Db 233 AIVVPGCIDPVQSTSMQ--YPHVLFDAQVPEPVIFTIPLRNSLYHL-MSDTTTSVLV 288
QY 180 AMLYTPI-----RNSGSDVFTVSCRILTRPSPEFDFYLVPPPTVESKTKPFTLPVLJT 233
Db 289 IMIYNDLINPYANDSNSSG-CIVTVE---TKPGDPFKFHLKPPG-----SMLT 333
QY 234 LGELNSRFPPLSIDEMVTSNESIVVQPNQGRVTLTGELLGTTLQACNICSRGKVTGQ 293
Db 334 HGS1-----PSDLI----- 342
QY 294 VPSQHMWNLEITNLNGTQDPTDDVPAPLGVPDFAGEVGV-----LSQRNRGESNPAN 348
Db 343 -PKSSSLM--IGNRHS--DITDFVIKPP-----VFQANRHFDFNQETAGWSTPRF 388
QY 349 RAHDAAVATYSKYTPKGLVQIGT-----MNTNDVENQPTKFTPIGLNEVANGH 398
Db 389 R---PITTVSEKSGSKLG-IGVATDSIVPGIPDGWPDTTI---PEKLTAGDYAITNGG 441
QY 399 RFEQWTLPRYSGALTLNWN 417
Db 442 NNDITTAADYDCASIIKKN 460

RESULT 3
US-10-670-695-4
; Sequence 4, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; TITLE OF INVENTION: STATUS OF AN ANIMAL
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-670-695-4

Query Match          10.3%; Score 292.5; DB 12; Length 547;
Best Local Similarity 22.7%; Pred. No. 1.4e-18;
Matches 133; Conservative 84; Mismatches 223; Indels 147; Gaps 25;

QY 11 SMDGAGLVPESSN--EAMALEPVGASLAAPVTGQNTIIPW-IRTNFVQAPNGEFTVS 67
```

```
Db      1  ADDGSI-TAPEOGTVMGVIAPSPSQMSTAADMATGSKVDSSEWAEFFSHTSNM-----S 55
Qy      68  PRNSPGEILVNLELGPENLPYLAHLARMYNGVAGGMEVQVMLAGNAGTAGKIIFAAVPPY 127
Db      56  TSETQKILFKQSLGPLNPLYLHAKLYAVWAGSIEVRFSGSGVGGKLAIVVPP- 114
Qy     128  FVENLSPSQITMFPVHIIDVRLTPVLLPMPDVRSTLPHFNQKDEPKMRLVAMLYTPLR 187
Db     115  -GVDPVQSTMLQYPHVLFDAQVPEVIFCLPDLASTLYHL-MSDTRDTTSLVIMVYNDL- 171
Qy     188  SNGSGDDVFTVSC--RILTRSPPEDEFTYLVP-----TVESKTKPFTLPLVLTGLGELS 238
Db     172  INPYANDANSSGCIYVETKPGDPFKHLLKPPGSMMLTHGSIPLSLIPKT-SSLWIG--- 227
Qy     239  NSRFLPSIDEMVTSP-----NE-----SIVVQPQNGR----- 265
Db     228  -NRYWSDITDFVIRPFVQANRHFDFNQETAGMSTPRPRPISVTITEQNGAKLGIGVATD 286
Qy     266  -----VTLDGELLGTTQLQACN-----ICSIKRGKVTQV-----P 295
Db     287  YIVPGIPDGPWDTTIPGELIPAGDYAITNGTGNDITATGYDTADIKNNTNFRGMYICG 346
Qy     296  SEQHWNLEITNLNGTQDPTDDVPAPLGVDPDFAGEVGVLSQRNGESNPANRAHDVAV 355
Db     347  SLQRAWGDK--KISNTAFITT-----ATLDGDNNNKINPCNTIDQSKI 387
Qy     356  ATYSKYTPKGLVQLGVTWNTNDVENQPTKFTPIGLNEVANGHREFQ-----WTLPRYSCA 411
Db     388  VVFQDNHVGGKA-----QTSDDTLALLGYT--GIGEAIGSDRDRVVRISTLPE-TGA 437
Qy     412  LTLNNLAPAVAPLPPGERLLPFRSYVPLKGGFNPADCSVPQBWQHFYQESAPSLGD 471
Db     438  R-----GGNHPIFYKNSIKL--GVVIRSIDV-----FNSQILHTSRQ 472
Qy     472  VALVRYVNPDTGRVLFKGLHGGFLTVSSTSTGVPVVPANGYFKPF 518
Db     473  LSLNHYLLPPDSFAVYRIIDNSGWFIDIGSDGFSFVGVSGFGKLE 519
```

RESULT 4

```
US-10-670-695-2
; Sequence 2, Application US/10670695
; Publication NO. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; STATUS OF AN ANIMAL
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; PRIOR FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 671
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-670-695-2
```

```
Query Match      10.3%; Score 292.5; DB 12; Length 671;
Best Local Similarity 22.7%; Pred. No. 28-18;
Matches 133; Conservative 84; Mismatches 223; Indels 147; Gaps 25;

Qy     11  SNDGAGLVPESNN--EAMALBPVVGASLAAPVTGNTIIDPW-IRTNFVQAPNGEFTVS 67
Db     125  ADDGSI-TAPEOGTVMGVIAPSPSQMSTAADMATGSKVDSSEWAEFFSHTSNM-----S 179
Qy     68  PRNSPGEILVNLELGPENLPYLAHLARMYNGVAGGMEVQVMLAGNAGTAGKIIFAAVPPY 127
```

```
Db     180  TSETQKILFKQSLGPLNPLYLHAKLYAVWAGSIEVRFSGSGVGGKLAIVVPP- 238
Qy     128  FVENLSPSQITMFPVHIIDVRLTPVLLPMPDVRSTLPHFNQKDEPKMRLVAMLYTPLR 187
Db     239  -GVDPVQSTMLQYPHVLFDAQVPEVIFCLPDLASTLYHL-MSDTRDTTSLVIMVYNDL- 295
Qy     188  SNGSGDDVFTVSC--RILTRSPPEDEFTYLVP-----TVESKTKPFTLPLVLTGLGELS 238
Db     236  INPYANDANSSGCIYVETKPGDPFKHLLKPPGSMMLTHGSIPLSLIPKT-SSLWIG--- 351
Qy     239  NSRFLPSIDEMVTSP-----NE-----SIVVQPQNGR----- 265
Db     352  -NRYWSDITDFVIRPFVQANRHFDFNQETAGMSTPRPRPISVTITEQNGAKLGIGVATD 410
Qy     266  -----VTLDGELLGTTQLQACN-----ICSIKRGKVTQV-----P 295
Db     411  YIVPGIPDGPWDTTIPGELIPAGDYAITNGTGNDITATGYDTADIKNNTNFRGMYICG 470
Qy     296  SEQHWNLEITNLNGTQDPTDDVPAPLGVDPDFAGEVGVLSQRNGESNPANRAHDVAV 355
Db     471  SLQRAWGDK--KISNTAFITT-----ATLDGDNNNKINPCNTIDQSKI 511
Qy     356  ATYSKYTPKGLVQLGVTWNTNDVENQPTKFTPIGLNEVANGHREFQ-----WTLPRYSCA 411
Db     512  VVFQDNHVGGKA-----QTSDDTLALLGYT--GIGEAIGSDRDRVVRISTLPE-TGA 561
Qy     412  LTLNNLAPAVAPLPPGERLLPFRSYVPLKGGFNPADCSVPQBWQHFYQESAPSLGD 471
Db     562  R-----GGNHPIFYKNSIKL--GVVIRSIDV-----FNSQILHTSRQ 596
Qy     472  VALVRYVNPDTGRVLFKGLHGGFLTVSSTSTGVPVVPANGYFKPF 518
Db     597  LSLNHYLLPPDSFAVYRIIDNSGWFIDIGSDGFSFVGVSGFGKLE 643
```

RESULT 5

```
US-10-209-507-2
; Sequence 2, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-2
```

```
Query Match      9.6%; Score 273; DB 14; Length 669;
Best Local Similarity 22.4%; Pred. No. 1,4e-16;
Matches 120; Conservative 74; Mismatches 188; Indels 154; Gaps 23;

Qy     67  SPRNSPGEILVNLELGPENLPYLAHLARMYNGVAGGMEVQVMLAGNAGTAGKIIFAAVPP 126
Db     180  STSETQKILFKQSLGPLNPLYLHAKLYAVWAGSVDVRFSGSGVGGKLAIVVPP 239
Qy     127  YFPVENLSPSQITMFPVHIIDVRLTPVLLPMPDVRSTLPHFNQKDEPKMRLVAMLYTPL 186
Db     240  --GVDPVQSTMLQYPHVLFDAQVPEVIFSIPLRSTLYHL-MSDTRDTTSLVIMVYNDL 296
```

[illegible]

Db 797 PEE-TSVPTETISTEVTVSPETTLTPTVPTVSTVTVNVSPEETSVPPEETILTLYT 855
QY 185 -----PLRNSGDDVFTVSCRILTRPSPEFDFYLVVPVTSKTKFTLVLTLGELSN 239
Db 856 EYVTPVTEVTVHTEVTVNVSPEETSVPTEETISTEVTVSPETTLTPTVPTVST-EVTN 914
QY 240 SRPPLSIDEMVTSNPSIV-----VOPQNGRVTLGELGTLTQLQACNICSIRGKVTGV 294
Db 915 ----VSPEETSVPPEETILTTEITVSPBETVPPIEGTLPTLEV-----TV 956
QY 295 PSEQHMMNLNLTNGTQDPTDDVPAPLGVPDFAGEVGV 335
Db 957 P-----IEVTF-----PTGETVPTVPTVSTEMTGV 984

RESULT 10
US-10-147-026-16
; Sequence 16, Application US/10147026
; Publication No. US20030003538A1
; GENERAL INFORMATION:
; APPLICANT: Dietrich, Paul Shartzer
; APPLICANT: Huang, Chiao-Chain
; APPLICANT: Johnson, Carl D.
; APPLICANT: Sangameswaran, Lakshim
; TITLE OF INVENTION: Neuropathic Pain Genes; Compositions
; TITLE OF INVENTION: Thereof; Related Reagents
; FILE REFERENCE: ROCH-006
; CURRENT APPLICATION NUMBER: US/10/147,026
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 60/155,702
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/189,931
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 6304
; TYPE: PRT
; ORGANISM: human
US-10-147-026-16

Query Match 3.9%; Score 111; DB 14; Length 6304;
Best Local Similarity 19.3%; Pred. No. 15;
Matches 100; Conservative 77; Mismatches 168; Indels 172; Gaps 25;

QY 42 TQGTNIIDPWIRTNVQA-----PNEGFTVSPRSPGEILVNLELGPENLYLAHARMY 96
Db 3884 TQPSVRRRGMEIAEIMTEENDDPRGIMFHVTRGAGEVITAYEVPPLNV----- 3934
QY 97 NGYAGMEVQVMLAGNAFTAGKIIFAAVPPYPPVENLSPSOLITWPHVVIDVRLTLEPVL 156
Db 3935 -----LQVPVRLAGSFGAVNYWKAGSDSAGLEDFFPS-----HGLLEFADKQVTAM 3982
QY 157 PMPDVRSTLTFHFNQDEPKRMLVAMLYTPLRSNGS-----GDDVFTVSCRILTRPSPEDF 212
Db 3983 -----IITII-----DDAEFELTEPTNTSLISVAGGRLGDV-----VVTWIPQND- 4026
QY 213 TVLVPTVTSKTKPFTLPLVTLTGELSNRPFLSIDEMVTS-----PNESI-----VVOPQNGRVT 267
Db 4027 -----SPF-----GVFGFEKTMIDESLSDDDPSVTILTVRSPGK----- 4065
QY 268 LDGELGTTLOACNICSIRGKVTGVQVPSQEHMMNLE-----ITNLNGT-QFDPTDDVP 320
Db 4066 -----GTVRL-----WTIDEKAKHNLSPNLGTLHFDETE----- 4095
QY 321 APLGVDPFAGEVGVLSQRNRGESNPANRAHDAVATYSDKTPKGLVQIGTWTNDVE 380
Db 4096 -----SQKT-----IVLHTLQTVLEEDRRFTLIQLISIDEVE 4127
QY 381 NQPTKFTP-----IGLNEVANG-----HRFEQWTLPRYSGALTLMNLAFAVAPLFFG 428
Db 4128 ISPVKGSASIIIRGDKKASGEVGIAPSSRILIGESAKYNGTAAIISLVRGPCIL-----G 4183

QY 429 ERLLFFRSYVPLKCGFGNPAIDCSVPOEWHQHFYQESAPSLGDLVALRVYVNPDP--TGRVL 486
Db 4184 EYTVFWRIFPPSVGSGFAETSKLTMRE-----QSAV-----IVVIQALNDDIPEEKSF 4232
QY 487 PEAKL-----HKGFLTVSSTSTGPPVVVPAN-----GYFKF 517
Db 4233 YEFQLTAVSEGGVLSSESSSTANITVWASDSPYGRFAF 4269

RESULT 11
US-10-207-706-2
; Sequence 2, Application US/10207706
; Publication No. US20030143681A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Anderson, Dirk W.
; TITLE OF INVENTION: Human Ataxin-1-Like Polypeptide IMX97018
; FILE REFERENCE: 3138-A
; CURRENT APPLICATION NUMBER: US/10/207,706
; CURRENT FILING DATE: 2002-07-26
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 689
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-706-2

Query Match 3.8%; Score 109.5; DB 14; Length 689;
Best Local Similarity 21.5%; Pred. No. 0.57;
Matches 120; Conservative 71; Mismatches 225; Indels 143; Gaps 30;

QY 5 SNDATPSNDG---AAGLVPSNEANEALPEVVGASLAAPVTGOTNIIDPWIRTNFVQAPN 61
Db 33 STNHTPSSDASESGVWVAGSQAGA-RVSLGDCGAEAITGLT--VDQYGMLYKVAVPP 89
QY 62 GEFTVSPRSPGEILVNLE-LGPELN---PYLAHARMYN--GYAGMEVQVMLAGNAFT 115
Db 90 ATF--SPTGLPS--VVMNSPLPPTFNVAASLIQHPGHYPLHYAQLPSTSLQFISGPS 145
QY 116 AGKIIFAAVPPYPPVENLSPS---QITMFPHVIDVRLTLEPVLPMDDVSTLPHENQKD 172
Db 146 ---LPYAVPPNPLPSPLLSPSANLATSHLPHFVFPYASLLAEGATFPPOAPSPAHSFNK-- 200
QY 173 EPKMLVAMLYTPLRSNGSGDDVFTVSCRILTRPSPEFDFYLVVPVTSKTKPFTL--- 229
Db 201 -----APSATSPSQ-----LPH--HSSTQPLDLAPG 225
QY 230 --PVLTLGELSNRPFLSIDEMVTS--NESIVOPQNGRVTLDGEL-----LGT 275
Db 226 RMPYI-----YQMSRLPAGYTLTHTETPPAGASPVLTPOESQSALEAAANGGORPRERNLVR 281
QY 276 TQLQACNICSIRGKVTGVQVPSQEHMMNLEITNLNGTQDPTDDVPAPL--GVPD----- 327
Db 282 RESEALDSPNSKGGQGLVPVVECVVVGQL--FSGSQ--TPRVEVAAPAHRTGPTDLDLVQ 338
QY 328 -----PAGEVFGVLSQRNRGESNPANRAHDAVATYSDKTPKGLVQIGTWTNDVENVQ 382
Db 339 RVVGALASQDYRVVAQRKEEPSPLNLSH-----HTP-----DHQSGR 377
QY 383 PTKFTPTIGLNEVANGHRF-----EQWTLPRYSGALTLMNLAFAVAPLFP-GER 430
Db 378 GSARNPAELAEKASQARGFYPSQHQEPVGRPLPK--AMVANGNLVPTGTDGSLLPVSGE 435
QY 431 LLP-----FRSYVPLKGGFGNPAIDCSVPOEWHQHFYQESAPSLGDLVALRVYVNPDTG 483
Db 436 ILVASSLDVQARATFPDKETPPPTTSSHLP-----SHFMKGAIILQATGELKRVEDLQTO 491
QY 484 RVLFEAKLHKGGFLTVSST 502
Db 492 DFVRSAEV-SGGLKIDSST 509

THIS PAGE BLANK (USPTO)

Result No.	Query Match	Score	Length	DB	ID	Description	
1	1872.5	65.8	539	1	COAT LORDV	P54635 lordsdale v	
2	1168.5	41.1	546	1	COAT SOUV3	P04542 southampton	
3	321.5	11.3	2344	1	POLN RHDV	P27410 rabbit hemo	
4	318	11.2	702	1	COAT SMSV1	P36284 san miguel	
5	296.5	10.4	2208	1	POLN MANCV	Q69014 manchester	
6	292.5	10.3	668	1	COAT FCVC6	P27404 feline cali	
7	292.5	10.3	671	1	COAT FCVC9	P27406 feline cali	
8	280.5	9.9	668	1	COAT FCVF4	P27405 feline cali	
9	274	9.6	703	1	COAT SMSV4	P36285 san miguel	
10	143	5.0	2193	1	POLG CX16T	Q9df31 c genome po	
11	138	4.9	2193	1	POLG CX16G	Q65900 c genome po	
12	138	4.9	2333	1	POLG FMDV1	P03306 f genome po	
13	137	4.8	2332	1	POLG FMDV0	P03305 f genome po	
14	136	4.8	1011	1	POLG FMDVT	P15072 foot-and-mo	
15	131.5	4.6	2153	1	POLG HRV16	Q82122 h genome po	
16	129	4.5	2336	1	POLG FMDVZ	P49303 f genome po	
17	127	4.5	2332	1	POLG FMDVA	P03308 f genome po	
18	124.5	4.4	2205	1	POLG POL2W	P23069 p genome po	
19	124	4.4	2194	1	POLG EC30B	Q9wn78 e genome po	
20	124	4.4	2194	1	POLG HE701	P32537 h genome po	
21	120.5	4.2	2207	1	POLG POL2L	P06210 p genome po	
22	118.5	4.2	2206	1	POLG POL32	P06209 poliiovirus	
23	118	4.1	2208	1	POLH POL1M	P03300 p genome po	
24	117.5	4.1	5376	1	ZAN MOUSE	O88799 mus musculu	
25	116	4.1	2432	1	Y43R IRV6	P18305 chilo iride	
26	115.5	4.1	2175	1	POLG BOVEV	P12915 b genome po	
27	114.5	4.0	2184	1	POLG EC01F	O91734 e genome po	
28	114.5	4.0	2193	1	POLG HE71B	Q66478 h genome po	
29	114	4.0	2206	1	POLG POL1M	P03299 p genome po	
30	113	4.0	2185	1	POLG SDVVU	P13900 s genome po	
31	113	4.0	2193	1	POLG HE71M	Q66479 h genome po	
32	112.5	4.0	2206	1	POLG POL3L	P03302 poliiovirus	
33	112	3.9	2209	1	POLG POL1S	P03301 p genome po	

```

RN  SEQUENCE FROM N.A.
RP  MEDLINE=93142023; PubMed=8380940;
RX  Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;
RT  "sequence and genome organization of a human small round-structured
RA  (Norwalk-like) virus.";
RL  Science 259:516-519(1993).
CC  -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation on its
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; L07418; AAA92984.1; -.
DR  PIR; B37491; B37491.
DR  InterPro; IPR004005; Calici_coat.
DR  InterPro; IPR008975; Viral_cap_coat.
DR  Pfam; PF00915; Calici_coat; 1.
KW  Coat protein; Glycoprotein.
FT  CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ  SEQUENCE 546 AA; 58774 MW; C82B2A85AD4B05EA CRC64;
Query Match 41.1%; Score 1168.5; DB 1; Length 546;
Best Local Similarity 44.9%; Pred No. 1.6e-78;

```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC EMBL; M57473; AAA47285.1; -
DR PIR; A41039; BRWHEB
DR PDB; 1KHV; 16-JAN-02.
DR MEROPS; C24.001; -
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR000317; Peptidase_C24.
DR InterPro; IPR000605; RNA helicase.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR001205; RNA pol P3D.
DR InterPro; IPR007094; RNA pol_PSVir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR Pfam; PF03510; Peptidase_C24; 1.
DR Pfam; PF00680; RNA dep RNA_pol; 1.
DR Pfam; PF00910; RNA helicase; 1.
DR PRINTS; PR00916; 2CENDOPTASE.
DR PRINTS; PR00918; CALICIVIRUSNS.
DR SMART; SM00382; AAA; 1.
KW Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;
KW Thiol protease; Helicase; ATP-binding; Coat protein; 3D-structure.
FT CHAIN 7508 7707 HELICASE (P2C-LIKE).
FT CHAIN 1109 1251 PROTEASE P3C.
FT CHAIN 7197 71625 PROTEASE P3C.
FT CHAIN ? 2344 RNA-DIRECTED RNA POLYMERASE.
FT CHAIN ? 2344 COAT PROTEIN.
FT NP_BIND 522 529 ATP (POTENTIAL).
FT ACT_SITE 1212 1212 PROTEASE (POTENTIAL).
FT ACT_SITE 1227 1227 PROTEASE (POTENTIAL).
SQ SEQUENCE 2344 AA; 257066 MW; 1454C248F81E9212 CRC64;

Query Match 11.3%; Score 321.5; DB 1; Length 2344;
Best Local Similarity 25.9%; Pred. No. 3.1e-15;
Matches 152; Conservative 79; Mismatches 208; Indels 147; Gaps 30;

QY 8 ATPNDGA-----AGLVPSNNEM-----ALPVPV-----GASLAAPVTG-----QTNIID 49
DB 1772 AAPOGEAAGTATTASVPGTTTDMGPGVVATTSVITAENSSAIATAGIGPQQVDOQE 1831

QY 50 PMRTNFVQAPNGEFTSPRNGPEILVNLELGPENLPYLAHLARMYNGYAGGMEVQVNL 109
DB 1832 TW-RNFYY--NDVFTWSVADAPGSLITYVQHSPPQNPPTAVLSQYAGWAGMGQFRFIV 1888

QY 110 AGNAFTAGKIIPAAVPPYPVENLSPS-QITMFPFHVIVDRTLEPVLPMDDVRSTLFHF 168
DB 1889 AGSGVFGRLVRAVTPPGI--EIGFGLVRFQPHVVDARSLEPVTITMPDLRPNMYH- 1944

QY 169 NOKDEPKM--RLVAMLYTDLRNGSGDDVFTVSCRILTRPSPFDFTYLVPP---TVESK 223
DB 1945 -PTGDPGLVPTLVLSYNNL--INPFGSTSAIQVTETRPSPDFEFVIRAPSSKTVDSI 2002

QY 224 TKP--FTLPLVTLGELSNRFLPSIDEMVTSNESIV-VOPQNGRT-----LDGELL 273
DB 2003 SPAGLLTTPVLTV-GVGNDRW-----NGQIVGLQVPVGGFSTCNRHNNLNGSTY 2050

QY 274 GTTQLOACNICIRGKVTQCVPESEQHMNLEITNLNGTQFDPTDDVPAPLGVPDF----- 328
DB 2051 GWSSPRFGDIDHRGASYSGSNATNLVQFWANAGSAIDNPISQV-APDGFPMDSFVPF 2109

QY 329 -----AGEV-FGVLSQRNGESNPAN-RAHDAAVAT-----YSDK 361
DB 2110 NGFGIPAAAGWVGAIWNSGAPNVTTVQAYELGFATGACGNLOPNTTTSQAQTVAKSI 2169

QY 362 YTPKGLVQ-----TGTWNTDVENQPT-----KFTPIGLN 392
DB 2170 YAVVTGTAQNPAGLFVMAAGIISTENASAITTPODRIVTTGTGTAAPVGNKTPIMPA 2229

QY 393 EV-----ANCHRPEQWTLPRYSGALTNMNLAPAVAPLFPGERLLPFRSYVP 439
DB 2230 SVVTRTGDVNTAGSANGTQGTGSQLP---PVTGLSLNNYSALMPGQ---FFWQOLT 2283

QY 440 LKGGFGNPAIDCSVPQWQHFYQESAPS-----LGDVALVRYNVP 480
DB 2284 PASGP--MEIGLSVD---GYFYAGTGASTLLIDTELIDVRPVG 2323

RESULT 4
COAT_SMSV1
ID COAT_SMSV1 STANDARD; PRT; 702 AA.
AC P36284;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE Coat protein (Capsid protein).
OS San Miguel sea lion virus (serotype 1) (SMSV 1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=36406;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92410750; PubMed=1529644;
RA Neill J.D.;
RT "Nucleotide sequence of the capsid protein gene of two serotypes of
RT San Miguel sea lion virus: identification of conserved and non-
RT conserved amino acid sequences among calicivirus capsid proteins.";
RL Virus Res. 24:211-222(1992).
CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

EMBL; M87481; AAA16217.1; -
DR PIR; A48562; A48562.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 493 493 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 702 AA; 77850 MW; E6E5A58523DEE3D7 CRC64;

Query Match 11.2%; Score 318; DB 1; Length 702;
Best Local Similarity 25.8%; Pred. No. 1e-15;
Matches 153; Conservative 74; Mismatches 199; Indels 168; Gaps 33;

QY 8 ATPNDGAAGLVPSNNEMALEPVPVVGAS-----LAAPVTGQTNIID-PMRTNFVQ---A 59
DB 151 AESDGPAGADIVTEBQGTVVQQPVPQAQSLTTLAAASTGKT--VDCW--TTFESYHTA 206

QY 60 PNGEFTVSPRNSPGBILVNLELGPENLPYLAHLARMYNGYAGGMEVQVNLNAGNATAGKI 119
DB 207 VNMSTT---EAQGKILFSLRSLSPENLPYLRHISLSYSTWSGGIDVTRTSGSGVFGGKL 262

QY 120 IFAAVPPYP-PVENLSPSQTMTFPHVIVDRTLEPVLPMDDVRSTLFHFNQKDPKQML 178
DB 263 AALIVPPGIEPVE--SPTML-QYPHVLPDARQTEPVTITPIDIKTLTH-SMDDTDTTRL 318

QY 179 VAMLYTPI-RLNSGDDVFTVSCRILTRPSPFDFTYLVPPPTVSKTKPFTLPLVTLGEL 237
DB 319 VIMVYNELINPVEQSEPKSSCSITVETRPSSDFTFSLKKPPG-----SLKKGSI 368

QY 238 SNRPPLSIDEMVTSNESIVVQPNQR-----VTLDGELLGTTTQLOACNICIRK 289

```

CC
DR EMBL: X86560; CRA60262.1; -.
DR MEROPS; C24.UEW; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR000317; Peptidase_C24.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_P5vir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam: PF00915; Calici_coat; 1.
DR
DR
DR

QY	62	G-----EFTVSPRNSPGEILVLMLEPELNPYLX 91
Db	1689	GAAQRLLEAVATGATQSNVPBAIRCFAVFTFAWDEMPGTGFLGSLSHPNINPYTSH 1748
QY	92	LARMYNGVAGMVEOVMLAGNAFTAGKIIFAAPVPYFFVEVENSQSQT---MPHFVILDV 148
Db	1749	LSGMWAGGSGSEVRLSISGSGVFAGRIIASVIPP-----GVDPSIRDPGVLPHAFVDA 1803
QY	149	RTLEPYLLPMPDVRLTHFNHOKDBPKMLVLMLYTPLRSNGSGDDVFTVSCRILTRSP 208
Db	1804	RITEPVSMIPDVRAVDYHRMDGABETCSLGFWYQPLNPFSTTAVSTCWVSVETKEGG 1863
QY	209	EDFTYVLVP---TVESKTKPFTLVLTGELSNSRFLPSIDEMWTSNPESIVVQPNQGRV 266
Db	1864	DFDFCLLRPFQOMENGVSPEGLLPRLUGYSGNRVGGVLVGMILVAESHKQVNRHFNS 1923
QY	267	TLDGELGTTQLOACNICSIKGVKTQVPS-BQHMWNLEITNLNQTQFDDTDVPAPLGV 325
Db	1924	VTFGNSTAPNPMWAEI-----VINQAHSTSRHAW-LSIGAQNKGPLFP-----GI 1968
QY	326	PDF-----AGEVEGVL-----SORNNGESNPA-----NRAHDAAVATY----- 358
Db	1969	PNHFPPDSCASTVVGAMDTSILGGRPSTGVCGPALSFQNGDGVYENDTFSVNFATYDPLTSG 2028
QY	359	-----SDKYTP-KLGLVQIG-----TW-----NTNDVENVQPT----- 384
Db	2029	TGVALTNSINPALSIALVRISNNDFTSGFANDKQVVVQMSWEMYCTNQIRGOVTPMSGT 2088
QY	385	--KFTPIGLNEVANGHRFEOWTLPRYSGALTLMNMLAPAVAPLFGERLLFFRYSVPLKG 442
Db	2089	NYFTSTGTANTLV-----LWQ-----ERMLSVD----- 2111
QY	443	GFGNPAIDCSPQFVQHFVOESAPSLGDVALVRYVNPDTGRVLFEAKLHKKGFLTVSST 502
Db	2112	--GHQAILYSSQLERTAEYFQNDIVNIPENSMAVF-NVETNSASFOGIRPDGYM-VTGG 2167
QY	503	STGPVVVPANGYFKFDSWVWQFSLAPMG-----TGN-CRRR 538
Db	2168	SIG-VNVPLEPETF-----OYGILPLSAAALSGPSPGNMGAR 2204

RESULT 6		
DATA	STANDARD;	PRT; 668 AA.
COAT FCVC6		
COAT FCVC6		
P27404:		
AG		
01-AUG-1992	(Rel. 23, Created)	
01-AUG-1992	(Rel. 23, Last sequence update)	
16-OCT-2001	(Rel. 40, Last annotation update)	
Coat protein	(capsid protein).	
DE		

OS Feline calicivirus (strain CFI/68 FIV) (FCV).
 OC Viruses; sRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OX NCBI_TaxID=11979;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=91374597; PubMed=1716692;
 RA Neill J.D., Reardon I.M., Heinrikson R.L.;
 RT "Nucleotide sequence and expression of the capsid protein gene of
 RT feline calicivirus";
 RL J. Virol. 65:5440-5447 (1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Neill J.D.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- PTM: The N-terminus is blocked.
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M32819; AAA42925.1; -;
 DR EMBL; U13992; AAC13993.1; -;
 DR PIR; A40507; VCMWF9.
 DR InterPro; IPR004005; Calici.coat.
 DR InterPro; IPR008975; Viral_Cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 668 AA; 73550 MW; 9E52312108D4D41 CRC64;
 Query Match 10.3%; Score 292.5; DB 1; Length 668;
 Best Local Similarity 26.0%; Pred. No. 7.3e-14;
 Matches 113; Conservative 53; Mismatches 143; Indels 125; Gaps 21;
 QY 11 SNDGAAGLVPSNN--EAMALEPVVGASLAAPVTQTNIIIDPW-IRTNFVOAPNGEFTVS 67
 DB 125 ADDGSI-TTPEQGTWGGVIAEPNAQMSAADMATKGSVDSEWEAFFSFHTSVNW----S 179
 QY 68 PRNSPGEILVNLELGPENLPYLAHARMYNGYAGGMEVQVLMAGNAFTAGKIIPAAPPY 127
 DB 180 TSETQKILFKQSLGFLNPLNLYTHLAKLYVAWSGSDVRFSGSGVFGGKLAATVVPFG 239
 QY 128 P-PVENLSPSQITMPEPHVIDVRLTLEPVLPMVDVRSITLPHFNQKDEPKMRVLVAMLYTPL 186
 DB 240 IDPVQSTSMQ---YPHFLDARQVEPVIFSPIDRLSTLYHL-MSDPTDTSILVIMVNDL 295
 QY 187 -----RNSGDDVFTVSC--RILTRPSPEFDTYLVPPVESKTPFTLPVLTGLSLS 240
 DB 296 INPYANDSNSSG-CIVTVE-----TKPGDFKHLKPG-----SMTHGSI--- 337
 QY 241 RPPLSIDEMVTSPNESIVVQPNQGRVTLTGELLGTTQLQACNICSGIRKGTQVPSQHM 300
 DB 338 -----PSDLI-----PKSSSL 348
 QY 301 WNLLETNLNGTQF--DPTDDVPAPLGVPDPFAGEVFGV-----LSQRNGESNPANRAHDA 353
 DB 349 W-----IGNRFWDITDFVIRPF-----VFQANRHFDFNQETAGWSTPRP---P 390
 QY 354 VVATYSDKYTPKLGVLQVIGT-----WNTNDVENQPTKFTPIGLNEVANGHFRFQW 403
 DB 391 ITITISVESAKLG-IGVATDYIVPGIDGWPDDTTI-----PGELVPVGDAYITNGTNDIT 446

QY 404 TLPFRYSGALTLNMN 417
 DB 447 TAAQYDAATEIRNN 460
 RESULT 7
 COAT_FCVF9
 ID COAT_FCVF9 STANDARD; PRT; 671 AA.
 AC P27406;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Coat protein (Capsid protein).
 CYP76.
 OS Feline calicivirus (strain F9) (FCV).
 OC Viruses; sRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OX NCBI_TaxID=11981;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92410623; PubMed=1529544;
 RA Carter M.J., Milton I.D., Meanger J., Bennett M., Gaskell R.M.,
 RA Turner P.C.;
 RT "The complete nucleotide sequence of a feline calicivirus.";
 RL Virology 190:443-448 (1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92339673; PubMed=1633955;
 RA Meanger J., Carter M.J., Gaskell R.M., Turner P.C.;
 RT "Cloning and sequence determination of the feline calicivirus strain
 RT F9.";
 RL Biochem. Soc. Trans. 20:265-268 (1992).
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M86379; AAA79327.1; -;
 DR EMBL; Z11536; CAA77636.1; -;
 DR PIR; B43382; VCMWF9.
 DR InterPro; IPR004005; Calici.coat.
 DR InterPro; IPR008975; Viral_Cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 618 618 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 671 AA; 73441 MW; 33BEE6D8370D5ES CRC64;
 Query Match 10.3%; Score 292.5; DB 1; Length 671;
 Best Local Similarity 22.7%; Pred. No. 7.3e-14;
 Matches 133; Conservative 84; Mismatches 223; Indels 147; Gaps 25;
 QY 11 SNDGAAGLVPSNN--EAMALEPVVGASLAAPVTQTNIIIDPW-IRTNFVOAPNGEFTVS 67
 DB 125 ADDGSI-TAPEQGTWGGVIAEPNAQMSAADMATKGSVDSEWEAFFSFHTSVNW----S 179
 QY 68 PRNSPGEILVNLELGPENLPYLAHARMYNGYAGGMEVQVLMAGNAFTAGKIIPAAPPY 127
 DB 180 TSETQKILFKQSLGFLNPLNLYTHLAKLYVAWSGIEVRFSGSGVFGGKLAATVVPFG 238
 QY 128 P-PVENLSPSQITMPEPHVIDVRLTLEPVLPMVDVRSITLPHFNQKDEPKMRVLVAMLYTPL 187
 DB 239 -GVDPVQSTSMQTPHVLFDARQVEPVIFCPLDRLSTLYHL-MSDPTDTSILVIMVNDL- 295
 QY 188 SNGSGDDVFTVSC--RILTRPSPEFDTYLVPP-----TVESKTKPFTLPVLTGLSLS 238

Matches 78; Conservative 36; Mismatches 93; Indels 37; Gaps 9;

QY 9 TFSNDAAGLVESNNEAMALEPVGASLAAPVGTQNIIDPW-IRTNFVQAPNGEFTVS 67
 132 TPEQGTAVG-----GVIAEPSAQNSTAADMASGKSVDSWEAFPSFHTSVNW---S 179
 QY 68 PRNSPGEILVLELPELNPYLHLARMVNGVAGMEVQVLMAGNAFTAGKIIFAAVPPY 127
 180 TSETQKILFKSLGGLPPLNPYLEHLSKLVVWSGSIYVRSISGSGVFGKLAIVWPP- 238
 QY 128 FVENLSPSQITMPPHVIIDVRLPVLPLMPDVRSTLPHFNQKDEPKMRLVAMLYTPL- 186
 239 -GVDPVQSTMLQYHVLFDARQVPEVIFTIPDLRSTLYHV-MSDTDTTSLVIMYNDLI 296
 QY 187 -----RNSGSGDDVFTVSCRIILTRSPDEFTYLPVPTVESKTKPFTLPVLTGLSLSNR 241
 297 NPYANDSNSG-CIVTVE-----TKPGDPFKFHLKPPG-----SVLTHGSIPLSD 341
 QY 242 FPLS 245
 342 IPKS 345

RESULT 9
 COAT_SMSV4
 ID COAT_SMSV4 STANDARD; PRT; 703 AA.
 AC P36285;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Coat protein (Capsid protein).
 OS San Miguel sea lion virus (serotype 4) (SMSV 4).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OX NCBI_TaxID=36407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92410750; PubMed=1529644;
 RA Neill J.D.;
 RT "Nucleotide sequence of the capsid protein gene of two serotypes of
 San Miguel sea lion virus: identification of conserved and non-
 conserved amino acid sequences among calicivirus capsid proteins.";
 RL Virus Res. 24:211-222(1992).
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M87482; AAA16220.1; -.
 DR PIR; C48562; C48562.
 DR InterPro; IPR004005; Calici coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 DR Coat protein; Glycoprotein.
 KW CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 482 482 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 703 AA; 77721 MW; C5DAD823B261073 CRC64;

Query Match 9.6%; Score 274; DB 1; Length 703;
 Best Local Similarity 28.4%; Pred. No. 1.8e-12;
 Matches 93; Conservative 54; Mismatches 142; Indels 38; Gaps 14;
 QY 8 ATSNDAAGLVESNNEAMALEPVGASLAAPVGTQNIIDPW-IRTNFVQAPNGE 63
 151 AESDGSAGLVESNNEAMALEPVGASLAAPVGTQNIIDPW-IRTNFVQAPNGE 206

Db 236 INPYANDANSSGCVITVETKPGDPKFKHLLKPPGSLTHGSIPLDKT-SSLWIG--- 351
 QY 239 NSRFLPSIDEMVTS-----NE-----SIVVQONGR----- 265
 Db 352 -NRYWSDITDFVIRFVFQANRHPDFNQETAGWSTPRFPISVITTEQNGAKLGIVATD 410
 QY 266 -----VTLDGELLGTTQLOACN-----ICSIRGKVTGV-----P 295
 Db 411 YIVGIPDGWPTTIPGELIPAGDVAITNGTNDITATGTDADIKNNTNFRGMYICG 470
 QY 296 SSOHWNLEITNLNGTQDFDTPDPAPLGVPDFAGEVFGVLSQRNRGSNPANRAHDAVV 355
 Db 471 SLQRAWGDK--KISNTAFITT-----ATLDGNNNNKINPCNTIDOSKI 511
 QY 356 ATYSOKYTPKGLVQIGTWNVDVENQPTKFTPIGLNEVANGHREQ-----WTLPRYSGA 411
 Db 512 VVFDQNHVGGKA-----QTSDDTLALLGYT--GIGEQAIGSDRDRVVRISTLPE-TGA 561
 QY 412 LTLNWNLAFAVAPLPPGERLLPFRSYVPLKGGFGNPAIDCSVPQEWQHFQESAPSLGD 471
 Db 562 R-----CGNHPIFYKSIKL--GVVIRSIDV-----FNSQILHTSRQ 596
 QY 472 VALVRYNPDGTGRVLPEAKLHGKGLFTVSTSTGPVVPVANGFYFKP 518
 Db 597 LSLNHYLLPDPSPAVYRIIDNSNGSWFDIGIDSGFSFGVSGFGKLE 643

RESULT 8
 COAT_FCVF4
 ID COAT_FCVF4 STANDARD; PRT; 668 AA.
 AC P27405;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Coat protein (Capsid protein).
 OS Feline calicivirus (strain Japanese F4) (FCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OX NCBI_TaxID=11980;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91306470; PubMed=1053578;
 RA Tohya Y., Taniguchi Y., Takahashi E., Utagawa E., Takeda N.,
 Miyamura K., Yamazaki S., Mikami T.;
 RT "Sequence analysis of the 3'-end of feline calicivirus genome.";
 RL Virology 183:810-814(1991).
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D90357; BAA14371.1; -.
 DR PIR; B40481; VCVWFC.
 DR InterPro; IPR004005; Calici coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 DR Coat protein; Glycoprotein.
 KW CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 668 AA; 73589 MW; 85BBD8B5804B503 CRC64;

Query Match 9.9%; Score 280.5; DB 1; Length 668;
 Best Local Similarity 32.0%; Pred. No. 5.6e-13;

Db 672 VNWDIDLNGYALRRKCELFYMRDAEFTFVAKNGELVPQLQYMYVPPGAPKPTSR 731
QY 301 -----WNLEITNLNGTQFDPTDDVPAPLQVP-----DEA 329
Db 732 DSFAWQ---TATNPSGVFKMTDPPAQVSPFMSPASAYQWFDYGYPTFGHQLANDLDYG 788
QY 330 -----GEVFGVLSORNG-ESNPANRAHDAVATYSKYTPKLGVLQIGTWTNDVNOPTK 385
Db 789 QCPNNMGTFSIRTGVTEKSP-----HSTLRVY-----WRIKHVRAWIPRLRNQPYL 837
QY 386 FTPIGLNEVANGHREFTQWTLPRYS 410
Db 838 FK-----INPNYKG 846

RESULT 11

POLG CX16G STANDARD; PRT; 2193 AA.
ID -POLG CX16G STANDARD; PRT; 2193 AA.
AC Q65900;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].
DE Cocksackievirus A16 (strain G-10).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.
OC NCBI_TaxID=69159;
OX [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=94303216; PubMed=8030260;
RA Poyry T., Hyypia T., Horsnell C., Kinnunen L., Hovi T., Stanway G.;
RT "Molecular analysis of coxsackievirus A16 reveals a new genetic group of enteroviruses."
RL Virology 202:982-987(1994).
CC -!- FUNCTION: It is thought that the P2C protein attaches to vesicular membranes and is associated with viral RNA synthesis.
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the poliovirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA) (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; U05876; AAA50478.1; --
CC HSSP; P03300; IPOV.
CC MEROPS; C03.022; --
CC MEROPS; C03.022; --
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR000199; Pept_3C_picorn.
CC InterPro; IPR000081; Peptidase_C3.
CC InterPro; IPR003136; Pico_P1A.
CC InterPro; IPR002527; Pico_P2B.
CC InterPro; IPR001676; Rhv.

InterPro; IPR000605; RNA_helicase.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR001205; RNA_pol_P3D.
InterPro; IPR007094; RNA_pol_Psvir.
InterPro; IPR008975; Viral_cap_coat.
Pfam; PF00548; Cys_protease-3C; 1.
Pfam; PF02226; Pico_P1A; 1.
Pfam; PF00947; Pico_P2A; 1.
Pfam; PF01552; Pico_P2B; 1.
Pfam; PF00073; rhv; 3.
Pfam; PF00680; RNA_dep_RNA_pol; 1.
Pfam; PF00910; RNA_helicase; 1.
PRINTS; PR00918; CALICVIRUSN.
ProDom; PD001125; Cys_protease_3C; 1.
ProDom; PD001306; Pico_P2A; 1.
ProDom; PD001274; Pico_P2B; 1.
SMART; SM00382; AAA; 1.
Polyprotein; Coat protein; Core protein; Transferase; RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate; Lipoprotein.
CHAIN 2 69 COAT PROTEIN VP4.
CHAIN 70 323 COAT PROTEIN VP2.
CHAIN 324 565 COAT PROTEIN VP3.
CHAIN 566 862 COAT PROTEIN VP1.
CHAIN 863 1012 CORE PROTEIN P2A.
CHAIN 1013 1111 CORE PROTEIN P2B.
CHAIN 1112 1440 CORE PROTEIN P2C.
CHAIN 1441 1526 CORE PROTEIN P3A.
CHAIN 1527 1548 GENOME-LINKED PROTEIN VP3.
CHAIN 1549 1731 PICORNAIN 3C.
CHAIN 1732 2193 N-myristoyl glycine (in host) (By similarity).
ACT SITE 1695 1695 PROTEASE (POTENTIAL).
ACT SITE 1709 1709 PROTEASE (POTENTIAL).
SEQUENCE 2193 AA; 243209 MW; 04B3BCE572A76E38 CRC64;
Query Match 4.9%; Score 138; DB 1; Length 2193;
Best Local Similarity 20.5%; Pred. No. 0.094;
Matches 101; Conservative 53; Mismatches 185; Indels 154; Gaps 21;
QY 44 QTNIDPWRITNF---VQAPNGEFTVSPRNSPGILNVNLELGPENLPYLAHLARMYNGYA 100
Db 382 QSNETTPQRLCFPVSVQSKTGELCAVFRADPG-----RNGPMQSTILGOLCRYTQWS 435
QY 101 GGMGVQVVMAGNAFTAGKLIIFAAVPP--YFVENLSPSQITMFPDHVIIDVRLTLEPVLPLM 158
Db 436 GSLVETFMFAGSFMATGKMLIAYTPPGGVPADRLTAMLGT---HVIWDFGLQSSVTLVI 492
QY 159 PDVRSTLPHFNKD-----EPKRLVAML-----YT----- 184
Db 493 PWISNTHPAHAKDGYFDYTTGTITWIYQTNVYVPIGAPTTAYIVALAAQDNFTMKLC 552
QY 185 -----PLNSGSGDDVFTVSCRILTRPSPEFDFTLVLPPTV-----ESTKPTFLPVL 232
Db 553 KOTEDIEQSANIQGGIADMIDQAVTSRGRALTSQVEPTAANTNASEHRLGTGLVPAL 612
QY 233 TLGEL---SNSRFPFLSID-----EMVTSFN-----ESTVQPOGNGRVTLDG-- 270
Db 613 QAAETGASSNAQENLLETRCVLNHHSIQETTIGNFFSRAGLVSITMTPTGTQNTDGYV 672
QY 271 -----ELIGTTQL-QACNICS-----IRKVTGQVPSE--QHM----- 300
Db 673 NWDIDLNGYAQMRKCELFYMRDAEFTFVAKNGELVPQLQYMYVPPGAPKPTSRD 732
QY 301 -----WNLEITNLNGTQFDPTDDVPAPLQVP-----DEA 329
Db 733 SFAMQ---TATNPSIFVKLTDPPAQVSPFMSPASAYQWFDYGYPTFGAHPQSNADADYQ 789
QY 346 -PANRAHDAVAVATYSKYTPK-----LGLVOIGTWTNDVNOPTKFTPIGLNEVANG 397
Db 790 CPNNMGTFSIRTGVTEKSPHSITLRYVRIKHVRAWIPRLRNQPYLKF----- 839

RESULT 14

POLG_FMDVT STANDARD; PRT: 1011 AA.

ID POLG_FMDVT STANDARD; PRT: 1011 AA.

AC P15072; Q84755; Q84756; Q84757; Q84758;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat

DE protein VP4; Coat protein VP2; Coat protein VP3; Coat protein VP1;

DE Core protein p12] (Fragment).

OS Foot-and-mouth disease virus (strain C1) (Aphthovirus C) (FMDV).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Aphthovirus.

OC NCBI_TaxID=12121;

OX NCBI_TaxID=12121;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94069803; PubMed=6316275;

RA Beck E., Fores S., Strebel K., Cattaneo R., Feil G.;

RT "Structure of the FMDV translation initiation site and of the

RL structural proteins.";

RL Nucleic Acids Res. 11:7873-7885(1983).

CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,

CC each of which is composed of one copy each of proteins VP1, VP2,

CC VP3, and VP4.

CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.

CC -!- MISCELLANEOUS: THE COAT PROTEIN VP1 CONTAINS THE MAIN ANTIGENIC

CC DETERMINANTS OF THE VIRION; THEREFORE, CHANGES IN ITS SEQUENCE

CC MUST BE RESPONSIBLE FOR THE HIGH ANTIGENIC VARIABILITY OF THE

CC VIRUS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

Nucleic Acids Res. 12:6587-6601(1984).

[2] RN
RN
RN SEQUENCE FROM N.A.
STRAIN=O1BFS;
MEDLINE=81143292; PubMed=6298715;
MAKOFF A.J., PAYNTER C.A., ROWLANDS D.J., BOOTHROYD J.C.;
"Comparison of the amino acid sequence of the major immunogen from
three serotypes of foot and mouth disease virus.";
Nucleic Acids Res. 10:8285-8295(1982).
[3] RN
RN X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
MEDLINE=89143740; PubMed=2537470;
ACHARYA R., FRY E., STUART D., FOX G., ROWLANDS D., BROWN F.;
"The three-dimensional structure of foot-and-mouth disease virus at
2.9-A resolution.";
Nature 337:709-716(1989).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
{RNA} (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
each of which is composed of one copy each of proteins VP1, VP2,
VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- MISCELLANEOUS: THE STRAIN GK SEQUENCE IS SHOWN
CC -!- MISCELLANEOUS: THE COAT PROTEIN VP1 CONTAINS THE MAIN ANTIGENIC
DETERMINANTS OF THE VIRION; THEREFORE, CHANGES IN ITS SEQUENCE
MUST BE RESPONSIBLE FOR THE HIGH ANTIGENIC VARIABILITY OF THE
VIRUS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; X00871; CAA25416.1; -;
EOL; J02185; AAA42635.1; -;
PDB; 1QMY; 18-SEP-01.
PDB; 1QQJ; 10-NOV-00.
MEKOPS; C03.008; -;
InterPro; IPR004004; Calici_pol_hel.
InterPro; IPR009003; Cys_Ser_cryptin.
InterPro; IPR004080; FMDVPicoat.
InterPro; IPR008739; Peptidase_C28.
InterPro; IPR001676; Rhv.
InterPro; IPR000505; RNA_helicase.
InterPro; IPR007095; RNA_pol_PS.
InterPro; IPR001205; RNA_pol_P3D.
InterPro; IPR007094; RNA_pol_Psvir.
InterPro; IPR008975; Viral_cap_coat.
Pfam; PF05408; Peptidase_C28; 1.
Pfam; PF00073; rhv; 3.
Pfam; PF06680; RNA_dep_RNA_pol; 1.
Pfam; PF00910; RNA_helicase; 1.
PRINTS; PR00918; CALICIVIRUSNS.
PRINTS; PR01542; FMDVP1COAT.
POLYPTEIN; Coat protein; Core protein; Core protein; RNA-directed RNA polymerase;
Transferase; Hydrolase; Thiol protease; Nonstructural protein;
Myristate; 3D-structure; Lipoprotein.
KW CHAIN 1 201 NONSTRUCTURAL PROTEIN P20A.
FT CHAIN 202 286 COAT PROTEIN VP4.
FT CHAIN 287 504 COAT PROTEIN VP2.
FT CHAIN 505 724 COAT PROTEIN VP3.
FT CHAIN 725 937 COAT PROTEIN VP1.
FT CHAIN 938 1107 CORE PROTEIN P12.
FT CHAIN 1108 1425 CORE PROTEIN P34.
FT CHAIN 1426 1578 CORE PROTEIN P14.
FT CHAIN 1426 1578 GENOME-LINKED PROTEIN VPg.
FT CHAIN 1579 1649 PROTEASE.
FT CHAIN 1650 1862 RNA-DIRECTED RNA POLYMERASE.
FT CHAIN 1863 2332 N-myrystoyl glycine (in host).
FT LIPID 202 202

```
CC EMBL; X00130; CAA24960.2; -.
DR PIR; A20288; GNYYCI.
DR HSP; Q88571; 1TME.
DR InterPro; IPR004080; FMDVP1coat.
DR InterPro; IPR008739; Peptidase_C28.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF05408; Peptidase_C28; 1.
DR Pfam; PF00073; rhv; 3.
DR PRINTS; PR01542; FMDVP1COAT.
KW Polyprotein; Coat protein; Core protein; Nonstructural protein;
KW Myristate; Lipoprotein.
FT CHAIN 1 216 NONSTRUCTURAL PROTEIN P20A.
FT CHAIN 217 286 COAT PROTEIN VP4.
FT CHAIN 287 504 COAT PROTEIN VP2.
FT CHAIN 505 723 COAT PROTEIN VP3.
FT CHAIN 724 932 COAT PROTEIN VP1.
FT CHAIN 933 >1011 CORE PROTEIN P12.
FT LIPID 217 217 N-myristoyl glycine (in host)
                                     (By similarity).
FT NON_TER 1011 1011
FT SEQUENCE 1011 AA; 111535 MW; 888845600A560601 CRC64;

Query Match      4.8%; Score 136; DB 1; Length 1011;
Best Local Similarity 20.5%; Pred. No. 0.044;
Matches 90; Conservative 52; Mismatches 143; Indels 155; Gaps 19;

QY 9 TPSNDGAAGLVPSNEAMALPVGASLAAPVTGQTNIIIDPWIRNFVQANGFTVSP 68
Db 263 TQNNWFSLA-----SFAFLGALLADKKTEILLERILT-----TRNGQTSTT 312
QY 69 RNSPG-----BILVNLELGPBLNPLYAH- 91
Db 313 QSSVGVTFGYATAEDSTSGPNTSGLETRVHQAEFFKALFDWPSQFNGHMKVLPHE 372
QY 92 -----LARMNGVAGMEVQVMLAGNAFTAGKIIFAAVPPYFPVENLSPSQITWPHV 144
Db 373 PRGVYGLVKSYAYRWNGWDVEVTAVGNQFNGCLQAALVPEMGDISREKYQLTLYPHQ 432
QY 145 IIDVRLTEPLVLPMPDVRSTLHFHVKQDEPKMR-----IVAMLYTPLRNSGDDVFTVSC 200
Db 433 FIPRNTNMTAHITVPV-----GVNRYDQYKQHRPWTLMVMVAPLTNTAGAQIKV-- 485
QY 201 RIILRPSPEFDLYVPVTVESKTPFTPLVLTLGELNSR--FPLSIDE-----WVTSNP 254
Db 486 -----YANTAPTNVH-----VAGELPSKEGIFPVACSDGYGNWVTTDP 523
QY 255 ESIVVQPQNGRV-----TLDGELLGTTOL-QAC-----NICSIRGKVTCQ----- 293
Db 524 KT--ADPAYGKVNPPTALPGRFTNYLDVAEACPTFLMFENVVYVSTRDQGRLLAKFD 581
QY 294 -VPSQHWNLEITNLNG--TQF-----DPTDD-----VPAPLGVPDFAGE 331
Db 582 VSLAAKHSNTYLAGLAQYTYTGTINLHFMTGPTDAKARYWAYVPPGMDAPD----- 637
QY 332 VFGVLSQRNGESNPANRAH 351
Db 638 -----NPEEAH 644

RESULT 15
POLG HRV16
ID POLG HRV16 STANDARD; PRT; 2153 AA.
AC Q82122;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2
DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein
DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-
DE linked protein VPG (P3B); Picornain 3C (SC 3.4.22.28) (Protease 3C)
DE (P3C); RNA-directed RNA polymerase P3D (SC 2.7.7.48)].
```

```
OS Human rhinovirus 16 (HRV-16).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Rhinovirus.
RX NCBI_TaxID=31708;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=95250310; PubMed=7732663;
RA Lee W.M., Wang W., Rueckert R.R.;
RT "Complete sequence of the RNA genome of human rhinovirus 16, a
RT clinically useful common cold virus belonging to the ICAM-1 receptor
RT group.";
RL Virus Genes 9:177-181(1995).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS) OF 2-853.
RX MEDLINE=94348864; PubMed=7915182;
RA Oliveira M.A., Zhao R., Lee W.M., Kremer M.J., Minor I.,
RA Rueckert R.R., Diana G.D., Pevear D.C., Dutko F.J., McKinlay M.A.,
RA Rossmann M.G.;
RT "The structure of human rhinovirus 16.";
RL Structure 1:51-68(1993).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 2-853, AND REVISIONS TO
RP 547-548.
RX MEDLINE=97238938; PubMed=9083115;
RA Hadfield A.T., Lee W.M., Zhao R., Oliveira M.A., Minor I.,
RA Rueckert R.R., Rossmann M.G.;
RT "The refined structure of human rhinovirus 16 at 2.15-A resolution:
RT implications for the viral life cycle.";
RL Structure 5:427-441(1997).
CC -|- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -|- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA) (N).
CC -|- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -|- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L24917; AAA69862.1; -.
DR PDB; 1AVN; 21-JAN-98.
DR PDB; 1AYM; 21-JAN-98.
DR MEROPS; C03.007; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR009003; Cys_ser_trypsin.
DR InterPro; IPR001199; Pept_3C_picorn.
DR InterPro; IPR000081; Peptidase_C3.
DR InterPro; IPR003138; Pico_F1A.
DR InterPro; IPR002527; Pico_P2B.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR006055; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_Fsvir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00548; Cys_vir_protease-3C; 1.
DR Pfam; PF02226; Pico_P1A; 1.
DR Pfam; PF00947; Pico_P2A; 1.
DR Pfam; PF03552; Pico_P2B; 1.
DR Pfam; PF00073; rhv; 3.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
```


DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PRO0918; CALICVIRUSNS.
DR ProDom; PD001125; Cys_protease_3C; 1.
DR ProDom; PD001306; Pico_P2A; 1.
DR ProDom; PD001274; Pico_P2B; 1.
DR SMART; SM00382; RAA; 1_
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
KW 3D-structure; Lipoprotein.
FT CHAIN 2 69
FT CHAIN 70 330 COAT PROTEIN VP4.
FT CHAIN 331 568 COAT PROTEIN VP2.
FT CHAIN 569 853 COAT PROTEIN VP3.
FT CHAIN 854 985 COAT PROTEIN VP1.
FT CHAIN 986 1090 CORE PROTEIN P2A.
FT CHAIN 1091 1412 CORE PROTEIN P2B.
FT CHAIN 1413 1489 CORE PROTEIN P2C.
FT CHAIN 1490 1510 CORE PROTEIN P3A.
FT CHAIN 1511 1693 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1694 2153 PICORNAIN 3C.
FT LIPID 2 2 RNA-DIRECTED RNA POLYMERASE P3D.
FT ACT_SITE 1657 1657 N-myristoyl glycine (in host).
FT ACT_SITE 1671 1671 PROTEASE (POTENTIAL).
FT CONFLICT 547 548 PROTEASE (POTENTIAL).
FT STRAND 27 30 KD -> NH (IN REF. 1).
FT HELIX 36 38
FT TURN 80 81
FT STRAND 83 87
FT TURN 88 89
FT STRAND 90 94
FT STRAND 101 102
FT HELIX 103 105
FT TURN 113 115
FT STRAND 123 123
FT TURN 126 128
FT HELIX 129 129
FT STRAND 133 134
FT STRAND 138 140
FT TURN 142 143
FT STRAND 147 151
FT TURN 152 152
FT HELIX 153 155
FT TURN 156 157
FT STRAND 159 167
FT STRAND 168 180
FT TURN 185 186
FT STRAND 188 197
FT TURN 198 198
FT STRAND 204 204
FT TURN 207 208
FT HELIX 213 216
FT STRAND 219 221
FT STRAND 223 224
FT TURN 225 227
FT TURN 232 233
FT STRAND 239 241
FT TURN 242 245
FT STRAND 246 246
FT HELIX 248 253
FT STRAND 256 260
FT TURN 261 263
FT STRAND 266 271
FT TURN 280 280
FT STRAND 282 284
FT TURN 285 285
FT STRAND 288 289
FT TURN 303 304
FT STRAND 307 323
FT TURN 338 341
FT STRAND 345 346
FT TURN 353 353
FT STRAND 356 357
FT STRAND 369 370

FT STRAND 372 372
FT TURN 373 373
FT HELIX 374 377
FT TURN 378 378
FT STRAND 381 382
FT TURN 385 386
FT STRAND 388 388
FT TURN 389 393
FT HELIX 398 398
FT STRAND 399 402
FT TURN 408 409
FT STRAND 411 416
FT TURN 419 420
FT TURN 422 426
FT TURN 428 433
FT TURN 434 435
FT STRAND 436 440
FT STRAND 443 449
FT TURN 453 454
FT STRAND 456 456
FT STRAND 458 464
FT HELIX 474 478
FT TURN 479 479
FT STRAND 481 486
FT STRAND 492 497
FT STRAND 506 507
FT TURN 513 514
FT STRAND 518 523
FT STRAND 528 528
FT TURN 531 532
FT STRAND 537 545
FT TURN 547 548
FT STRAND 550 554
FT STRAND 567 567
FT HELIX 570 579
FT TURN 580 580
FT STRAND 584 584
FT STRAND 588 588
FT TURN 600 601
FT STRAND 602 603
FT HELIX 605 607
FT HELIX 615 618
FT STRAND 624 624
FT STRAND 629 629

Query Match 4.6%; Score 131.5; DB 1; Length 2153;
Best Local Similarity 20.8%; Pred. No. 0.28; Indels 173; Gaps 24;
Matches 108; Conservative 61; Mismatches 177;
QY 48 IDPWRTNFWQAPNGE-----FTVSPRNSPGE-----ILVNLELGPENLPYLAHLARMYNG 98
Db 379 VDTLPIINSTOSNIGNVSMYTVTSLSPQTKLAEIEFAIKVDIASHPLATTLIGEIASYFTH 438
QY 99 YAGMEVQVMLAGNAFTAGKIIFAAVPPYFPVENLSPSQITMFPFHVIVDVITLSPVLLPM 158
Db 439 WTGSLRPFSEMFCGTANTTLKULLAVTPPGIG-KPRSRKEAMLGTHVWVMDVGLQSTVSLVV 497
QY 159 PDVRSTLTFHFNOKDEPKMRLVAMLYTPLRNSGSDVDFTVSCRILTRSPSPFDFTYLYPP 218
Db 498 PWISASQYRFT-----TPDTYSAG-----YITCWYQT-----NFVVPP 531
QY 219 -----TVE-----SKTKPFTLPVLTIGELSNRPPLS-----IDEMVTSNIESIVV--- 259
Db 532 NTNTAEMLCFVSGCKDFCLRMARDTDLHKQTGPITQNPVERVYDDEVL---NEVLVWPN 588
QY 260 -----QPQNGRVTLDGELLG-TTQLOACNICIRGKVTGOVSEQHMMNLEITNLNGTQF 313
Db 589 NQSHPTTSNAAAPVLDAAEATGHTNKIQPEDTITRYVQSSQTLDE----- 632
QY 314 DPTDDVPAPLGVDFEAGEVGVLSQRNRCESNPANRAHDV---VATYSDKYTPKLGVLVQ 370
Db 633 -----MSVESFLGR-----SGCIESVLIDVNDYNDQSFTK----- 663

QY 371 IGTWNTDVENQTKETPIGLNEVANGHR-FEQWTLPRYSGALTLMNMLAPAVP----- 424
Db 664 -----WNIN-----LOEMAQIRKFKFEMFTYAREDSIT-----MVPSVAAKDCHI 703
QY 425 -----LFGERLLFFERSYVPLKGGFNPDAIDCSVPQEWQHFYQESAPSL----- 469
Db 704 GHIVMQYMYVPGAPIPTTRDDYAWQSG-----TNASVFQHQPPPPRESLPLSLSTASAY 758
QY 470 -----GDVALVRY---VNPDTG-----RVLFEAKLHK 493
Db 759 YMFYDGYDGTYSRYGTVTNDMGTLCSRIVTSEQLHK 797

Search completed: June 1, 2004, 13:47:20
Job time : 8.52349 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 45.9597 Seconds
(without alignments)
3289.030 Million cell updates/sec

Title: US-09-926-799-8
Perfect score: 2843
Sequence: 1 MKMASNDAPSDGAGLVP.....VNQFYSLAPMTGNGRRRVQ 535

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2843	100.0	535	4 AAB49707	Aab49707 Small rou
2	2269.5	79.8	540	4 AAB49706	Aab49706 Small rou
3	2117.5	74.5	548	4 AAB49705	Aab49705 Small rou
4	2069.5	72.8	548	5 AAU91272	Aau91272 Norwalk v
5	2062.5	72.5	550	4 AAB49709	Aab49709 Small rou
6	2048.5	72.1	542	4 AAB49708	Aab49708 Small rou
7	1919	67.5	539	4 AAB49704	Aab49704 Small rou
8	1656	58.2	541	4 AAB49710	Aab49710 Small rou
9	1200.5	42.2	530	4 AAB49701	Aab49701 Small rou
10	1196.5	42.1	530	2 AAR50972	Aar50972 Norwalk v
11	1196.5	42.1	530	7 ADC72176	Adc72176 Norwalk v
12	1175	41.3	546	4 AAB49702	Aab49702 Small rou
13	1174.5	41.3	545	4 AAB49700	Aab49700 Small rou
14	1151.5	40.5	530	2 AAR57091	Aar57091 Small rou
15	1145	40.3	544	4 AAB49703	Aab49703 Small rou
16	320.5	11.3	579	2 AAW08143	Aaw08143 RHDV caps
17	231	10.2	622	4 AAB47045	Aab47045 Feline ca
18	230	10.2	623	4 AAB47044	Aab47044 Feline ca
19	286	10.1	623	4 AAB47043	Aab47043 Feline ca
20	284	10.0	668	4 AAB67462	Aab67462 Amino aci
21	277.5	9.8	547	4 AAM50108	Aam50108 Feline ca
22	277.5	9.8	669	4 AAB67461	Aab67461 Amino aci
23	277.5	9.8	671	4 AAM50107	Aam50107 Feline ca
24	272.5	9.6	668	4 AAR10686	Aar10686 Feline ca
25	272.5	9.6	668	4 AAE04304	Aae04304 Feline ca

26	194	6.8	40	5 AAU91274	Aau91274 Norwalk v
27	174	6.1	40	5 AAU91273	Aau91273 Norwalk v
28	130	4.6	934	1 AAP20016	Aap20016 Sequence
29	125.5	4.4	1147	5 ABB76724	Abb76724 Foot and
30	125.5	4.4	4643	4 ABB71609	Abb71609 Drosophil
31	122	4.3	980	2 AAW30611	Aaw30611 Arabidops
32	122	4.3	980	5 ABB91612	Abb91612 Herbicida
33	122	4.3	985	2 AAW30607	Aaw30607 Arabidops
34	121	4.3	761	2 AAR70358	Aar70358 Human NF-
35	117	4.1	338	7 ADC06674	Adc06674 Nervous n
36	114.5	4.0	366	5 ABB05228	Abb05228 Redspotte
37	111	3.9	930	6 ABP96278	Abp96278 Human NF-
38	111	3.9	2206	2 AAR22210	Aar22210 True type
39	110.5	3.9	40	5 AAU91275	Aau91275 Norwalk v
40	110.5	3.9	2468	6 ABU38411	Abu38411 Protein e
41	110.5	3.9	2468	6 ABP59933	Abp59933 Microbial
42	110	3.9	716	2 AAR70357	Aar70357 NF-AT tra
43	110	3.9	716	2 AAW02249	Aaw02249 Human tra
44	110	3.9	716	3 AAY96565	Aay96565 Human nuc
45	110	3.9	716	3 AAY96535	Aay96535 Human nuc

ALIGNMENTS

RESULT 1
AAB49707
ID AAB49707 standard; protein; 535 AA.
XX
AC AAB49707;
XX
DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 8.
XX
KW Small round structured virus; SRSV; food poisoning.
XX
OS Small round structured virus.
XX
PN WO200079280-A1.
XX
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000MO-JP004095.
XX
PR 22-JUN-1999; 99JP-00175928.
XX
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
WPI: 2001-080848/09.
DR N-PSDB; AAF29148.
XX
PT Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.
XX
PS Claim 1; Page 57-59; 84pp; Japanese.
CC This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks
SQ Sequence 535 AA;
Query Match 100.0%; Score 2843; DB 4; Length 535;
Best Local Similarity 100.0%; Pred. No. 2.2e-252;

Matches 535; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVANSNDAAAPSDNGAAGLVPEANNETMALEPVAGASIAAPLTGONNIIDPWIRLNFVQAP 60

Db 1 MKVANSNDAAAPSDNGAAGLVPEANNETMALEPVAGASIAAPLTGONNIIDPWIRLNFVQAP 60

QY 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLARMYNGYAGGVEVQVLLAGNAFTAGKLV 120

Db 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLARMYNGYAGGVEVQVLLAGNAFTAGKLV 120

QY 121 FFAVPPHFFPLENISPGQITMFPVHVIDVRLTPEVLLPLPDVVRNFFHYNQNEPRMRLVA 180

Db 121 FFAVPPHFFPLENISPGQITMFPVHVIDVRLTPEVLLPLPDVVRNFFHYNQNEPRMRLVA 180

QY 181 MLYTPLRSNGSGDDVFTVSCRVLTRSPDFDNLYVPPTLESKTKPFTLPILTIGELTNS 240

Db 181 MLYTPLRSNGSGDDVFTVSCRVLTRSPDFDNLYVPPTLESKTKPFTLPILTIGELTNS 240

QY 241 RFPVPIDELYTSFNESLVVQPNQRCALDGLQGTTLQLLPTAICSPFRGRINQKVSGENHV 300

Db 241 RFPVPIDELYTSFNESLVVQPNQRCALDGLQGTTLQLLPTAICSPFRGRINQKVSGENHV 300

QY 301 WNMQVNINGTPTPDTGDPVAPLGTDFSGKLFVLSQRDHONACKSHDAVIATNSAKFT 360

Db 301 WNMQVNINGTPTPDTGDPVAPLGTDFSGKLFVLSQRDHONACKSHDAVIATNSAKFT 360

QY 361 PKLGAIOIGTWEEDDVHINOPTKFTPVGLFENEGFNQWTLPNYSGALTLNMGALPPVAPT 420

Db 361 PKLGAIOIGTWEEDDVHINOPTKFTPVGLFENEGFNQWTLPNYSGALTLNMGALPPVAPT 420

QY 421 FPGEQILFFRSHIPKGGVADPVIDCILLPOEWIQLHYQESAPQSQSDVALIRFTNPTGRV 480

Db 421 FPGEQILFFRSHIPKGGVADPVIDCILLPOEWIQLHYQESAPQSQSDVALIRFTNPTGRV 480

QY 481 LFEAKLHRSYITVANTGSRPIVVPVANGYFRFDTWVNFYSLAPMGTCNGRRRVQ 535

Db 481 LFEAKLHRSYITVANTGSRPIVVPVANGYFRFDTWVNFYSLAPMGTCNGRRRVQ 535

RESULT 2

AAB49706

ID AAB49706 standard; protein; 540 AA.

XX AC AAB49706;

XX DT 04-APR-2001 (first entry)

XX DE Small round structured virus protein SEQ ID 7.

XX KW Small round structured virus; SRSV; food poisoning.

XX OS Small round structured virus.

XX PN WO200079280-A1.

XX PD 28-DEC-2000.

XX PF 22-JUN-2000; 2000WO-JP004095.

XX PR 22-JUN-1999; 99JP-00175928.

XX PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

XX PA (DENK-) DENKA SEIKEN KK.

XX PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

XX DR WPI: 2001-080848/09.

XX DR N-PSDB; AAP29147.

XX PT Kit for the detection and typing of small round-structured virus (SRSV)

XX PT strains for investigation of food poisoning outbreaks, contains

XX PT antibodies.

PS Claim 1; Page 54-57; 84pp; Japanese.

XX This invention relates to a kit for the detection and typing of small

CC round structured virus (SRSV) strains. The kit contains antibodies

CC directed against peptides represented in sequences AAB49700 - AAB49710,

CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -

CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is

CC used for detecting and typing strains of SRSV in order to prevent the

CC spread of infection and to examine the epidemiology of outbreaks

XX SQ Sequence 540 AA;

Query Match 79.8%; Score 2269.5; DB 4; Length 540;

Best Local Similarity 76.3%; Pred. No. 1.4e-199;

Matches 413; Conservative 65; Mismatches 56; Indels 7; Gaps 3;

QY 1 MKVANSNDAAAPSDNGAAGLVPEANNETMALEPVAGASIAAPLTGONNIIDPWIRLNFVQAP 60

Db 1 MKVANSNDAAAPSDNGAAGLVPEANNETMALEPVAGASIAAPLTGONNIIDPWIRLNFVQAP 60

QY 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLARMYNGYAGGVEVQVLLAGNAFTAGKLV 120

Db 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLARMYNGYAGGVEVQVLLAGNAFTAGKLV 120

QY 121 FFAVPPHFFPLENISPGQITMFPVHVIDVRLTPEVLLPLPDVVRNFFHYNQNEPRMRLVA 180

Db 121 FFAVPPHFFPLENISPGQITMFPVHVIDVRLTPEVLLPLPDVVRNFFHYNQNEPRMRLVA 180

QY 181 MLYTPLRSNGSGDDVFTVSCRVLTRSPDFDNLYVPPTLESKTKPFTLPILTIGELTNS 240

Db 181 MLYTPLRSNGSGDDVFTVSCRVLTRSPDFDNLYVPPTLESKTKPFTLPILTIGELTNS 240

QY 241 RFPVPIDELYTSFNESLVVQPNQRCALDGLQGTTLQLLPTAICSPFRGRINQKVSGENHV 300

Db 241 RFPVPIDELYTSFNESLVVQPNQRCALDGLQGTTLQLLPTAICSPFRGRINQKVSGENHV 300

QY 301 WNMQVNINGTPTPDTGDPVAPLGTDFSGKLFVLSQRDHONACKSHDAVIATNSA 357

Db 301 WNMQVNINGTPTPDTGDPVAPLGTDFSGKLFVLSQRDHONACKSHDAVIATNSA 357

QY 358 KFTPKLGAIOIGTWEEDDVHINOPTKFTPVGLFENEGFNQWTLPNYSGALTLNMGAL 414

Db 358 KFTPKLGAIOIGTWEEDDVHINOPTKFTPVGLFENEGFNQWTLPNYSGALTLNMGAL 414

QY 361 KYTPKGLVQLGTWNTNDEVE-NQTKFTPIGLNEVANGHREQWTLPRYSGALTLNWL 419

Db 361 KYTPKGLVQLGTWNTNDEVE-NQTKFTPIGLNEVANGHREQWTLPRYSGALTLNWL 419

QY 415 PPAVPTFECQILFFRSHIPKGGVADPVIDCILLPOEWIQLHYQESAPQSQSDVALIRFTN 474

Db 415 PPAVPTFECQILFFRSHIPKGGVADPVIDCILLPOEWIQLHYQESAPQSQSDVALIRFTN 474

QY 420 PPAVPLFFGERLLFPFRSYVPLKGGFGNPAIDCSVPQEWVQHYQESAPSLGDVALRYVN 479

Db 420 PPAVPLFFGERLLFPFRSYVPLKGGFGNPAIDCSVPQEWVQHYQESAPSLGDVALRYVN 479

QY 475 PDTGRVLFPEAKLHRSYITVANTGSRPIVVPVANGYFRFDTWVNFYSLAPMGTCNGRRRV 534

Db 475 PDTGRVLFPEAKLHRSYITVANTGSRPIVVPVANGYFRFDTWVNFYSLAPMGTCNGRRRV 534

QY 480 PDTGRVLFPEAKLHRSYITVANTGSRPIVVPVANGYFRFDTWVNFYSLAPMGTCNGRRRV 539

Db 480 PDTGRVLFPEAKLHRSYITVANTGSRPIVVPVANGYFRFDTWVNFYSLAPMGTCNGRRRV 539

QY 535 Q 535

Db 540 Q 540

RESULT 3

AAB49705

ID AAB49705 standard; protein; 548 AA.

XX AC AAB49705;

XX DT 04-APR-2001 (first entry)

XX DE Small round structured virus protein SEQ ID 6.

XX KW Small round structured virus; SRSV; food poisoning.

XX OS Small round structured virus.

XX PN WO200079280-A1.

PD 28-DEC-2000.
 XX
 PF 22-JUN-2000; 2000WO-JP004095.
 XX
 PR 22-JUN-1999; 99JP-00175928.
 XX
 XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.
 XX
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 XX
 XX WPI: 2001-080848/09.
 DR N-PSDB; AAP29146.
 DR
 XX
 PT Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX
 XX Claim 1; Page 52-54; 84pp; Japanese.
 PS
 CC This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAP20141 -
 CC AAP20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX
 XX Sequence 548 AA;
 SQ
 Query Match 74.5%; Score 2117.5; DB 4; Length 548;
 Best Local Similarity 70.5%; Pred. No. 1.4e-185;
 Matches 387; Conservative 74; Mismatches 73; Indels 15; Gaps 4;
 QY 1 MKMASNDRAAPSDGAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
 DB 1 MKMASNDRAAPSDGAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
 QY 61 NGFEVTSRNSPGEVLLNLELGPENLPYLAHLARMYNGYAGGVEVQVLLAGNAFTAGKLV 120
 DB 61 GGEFTVSPNSPGEVLLNLELGPENLPYLAHLARMYNGYAGGVEVQVLLAGNAFTAGKII 120
 QY 121 FAAPVPHPLENISPGQITMFPVHVIDVRLPVLPLDVRNPFHYNQNEPRMLVA 180
 DB 121 FAAPVPHPLENISPGQITMFPVHVIDVRLPVLPLDVRNPFHYNQNEPRMLVA 180
 QY 181 MLYTLRNSGDDVFTVSCRVLTRPSDFDFNVLVPPTLESKTPFTLPILTIGELTNS 240
 DB 181 MLYTLRNSGDDVFTVSCRVLTRPSDFDFNVLVPPTLESKTPFTLPILTIGELTNS 240
 QY 241 RPPVPIDELTSPNESLVVQONGRCALDGELOTTQLLPTAICSPGRINOKVSGE--- 297
 DB 241 RPPVPIDELTSPNESLVVQONGRCALDGELOTTQLLPTAICSPGRINOKVSGE--- 297
 QY 301 ADTATPRLFNYYWHVQDLNLTGTPYDPAEDIPGLTDPFRGKFGVASQORNDSTTRAH 360
 DB 301 ADTATPRLFNYYWHVQDLNLTGTPYDPAEDIPGLTDPFRGKFGVASQORNDSTTRAH 360
 QY 349 DAVIATNSAKFTPKLCAIGITWEEDVHINQKTPVGL-FENEG-FNQTWLPNYSGA 406
 DB 349 DAVIATNSAKFTPKLCAIGITWEEDVHINQKTPVGL-FENEG-FNQTWLPNYSGA 406
 QY 361 EAKVDITAGRTFPKLGSLISIT-DSDDFDQONQKFTPTVGVIGDNEAFQOWSLDYSQQ 419
 DB 361 EAKVDITAGRTFPKLGSLISIT-DSDDFDQONQKFTPTVGVIGDNEAFQOWSLDYSQQ 419
 QY 407 LTLNMLGAPVAPTPGEOILFRSHIPLKGGVADPVIDCLLPQEWIQLHYOESAPSQSD 466
 DB 407 LTLNMLGAPVAPTPGEOILFRSHIPLKGGVADPVIDCLLPQEWIQLHYOESAPSQSD 466
 QY 420 FTHNMMLAFAVAPNPFGEQQLLFRSOLPSSGSRNGVLDCLVPQEWQHFYQESAPATQ 479
 DB 420 FTHNMMLAFAVAPNPFGEQQLLFRSOLPSSGSRNGVLDCLVPQEWQHFYQESAPATQ 479
 QY 467 VALIRFTNPDTRVLPEAKLHRSYITVANTGSRPIVWPANGYFRPDTWVNOFYSLAPMG 526
 DB 467 VALIRFTNPDTRVLPEAKLHRSYITVANTGSRPIVWPANGYFRPDTWVNOFYSLAPMG 526
 QY 480 VALVRYVNDTGVKLVPEAKLHKLGFMTIANNGSDSPITVPPNGYFRFESWVNFYTLAPMG 539
 DB 480 VALVRYVNDTGVKLVPEAKLHKLGFMTIANNGSDSPITVPPNGYFRFESWVNFYTLAPMG 539
 QY 527 TNGRRRVQ 535
 DB 527 TNGRRRVQ 535
 QY 540 TNGRRRIQ 548
 DB 540 TNGRRRIQ 548

RESULT 4

AAU91272

ID AAU91272 standard; protein; 548 AA.

XX AC AAU91272;

XX DT 18-JUN-2002 (first entry)

XX DE Norwalk virus associated polynucleotide #1.

XX KW Norwalk virus; monoclonal antibody; geno group I; geno group II;

XX KW immunological detection; food; viral infection.

XX OS Norwalk virus.

XX PN JP2002020399-A.

XX PD 23-JAN-2002.

XX PF 10-JUL-2000; 2000JP-00208151.

XX PR 10-JUL-2000; 2000JP-00208151.

XX PA (OSAP) OSAKA PREFECTURE.

XX PA (NISE-) NIPPON SEIBUTSU KAGAKU CENT KK.

XX PA (IATR) IATRON LAB INC.

XX DR WPI: 2002-287412/33.

XX XX A monoclonal antibody useful in the immunological detection and diagnosis

XX PT of Norwalk virus infection.

XX PS Disclosure; Page 12-13; 24pp; Japanese.

XX CC The invention describes a monoclonal antibody recognising Norwalk virus,

XX CC a capsid protein of Norwalk virus, or a common antigen epitope on the

XX CC capsid protein molecule of geno group I and geno group II. The antibody

XX CC is useful for immunological detection and quantitative analysis of

XX CC Norwalk virus in foods and the serum of infected patients. This sequence

XX CC represents a Norwalk virus associated protein described in the invention

XX SQ Sequence 548 AA;

Query Match 72.8%; Score 2069.5; DB 5; Length 548;

Best Local Similarity 69.7%; Pred. No. 3.7e-181;

Matches 380; Conservative 73; Mismatches 77; Indels 15; Gaps 4;

QY 1 MKMASNDRAAPSDGAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60

DB 1 MKMASNDRAAPSDGAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60

QY 61 NGFEVTSRNSPGEVLLNLELGPENLPYLAHLARMYNGYAGGVEVQVLLAGNAFTAGKLV 120

DB 61 GGEFTVSPNSPGEVLLNLELGPENLPYLAHLARMYNGYAGGVEVQVLLAGNAFTAGKII 120

QY 121 FAAPVPHPLENISPGQITMFPVHVIDVRLPVLPLDVRNPFHYNQNEPRMLVA 180

DB 121 FAAPVPHPLENISPGQITMFPVHVIDVRLPVLPLDVRNPFHYNQNEPRMLVA 180

QY 181 MLYTLRNSGDDVFTVSCRVLTRPSDFDFNVLVPPTLESKTPFTLPILTIGELTNS 240

DB 181 MLYTLRNSGDDVFTVSCRVLTRPSDFDFNVLVPPTLESKTPFTLPILTIGELTNS 240

QY 241 RPPVPIDELTSPNESLVVQONGRCALDGELOTTQLLPTAICSPGRINOKVSGE--- 297

DB 241 RPPVPIDELTSPNESLVVQONGRCALDGELOTTQLLPTAICSPGRINOKVSGE--- 297

QY 298 -----NHVNMVQNTINGTDPDTPGDPAPLGTDFSGKLGVLQORDHNCARSH 348

DB 301 ADIATPRLFNYYWHVQDLNLTGTPYDPAEDIPGLTDPFRGKFGVASQORNDSTTRAH 360

CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX
 XX
 SQ Sequence 542 AA;

Query Match 72.1%; Score 2048.5; DB 4; Length 542;
 Best Local Similarity 71.7%; Pred. No. 3.1e-179;
 Matches 392; Conservative 64; Mismatches 74; Indels 17; Gaps 7;
 QY 1 MKMASNDAPSDNDGAGLVPEANNTMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
 DB 1 MKMASNDAPSDNDGAGLVPEANNTMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
 QY 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLARMYNGYAGVEVQVLLAGNAFTAGKLV 120
 DB 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLARMYNGYAGVEVQVLLAGNAFTAGKLV 120
 QY 121 FAAPVPHPLENISPGQITMFPVHVIDVRLTLEPVLLPDVNRNFFHYNQNEPRMLVA 180
 DB 121 FAAPVPHPLENISPGQITMFPVHVIDVRLTLEPVLLPDVNRNFFHYNQNEPRMLVA 180
 QY 181 MLYTPLRNSGDDVFTVSCRVLTSPDPPFNLYPPPTLESKTKPFTLPILTIGELTNS 240
 DB 181 MLYTPLRNSGDDVFTVSCRVLTSPDPPFNLYPPPTLESKTKPFTLPILTIGELTNS 240
 QY 241 RFPVPIDELTSPNESLVVQPNQRCALDGELOQTQLPTAICFRGRINQKVSNGHV 300
 DB 241 RFPVPIDELTSPNESLVVQPNQRCALDGELOQTQLPTAICFRGRINQKVSNGHV 300
 QY 301 WN---MQVTNINGTPTDPTGDPVAPLGTDFSGKLFGLVLSQRDHDN-----ACRSHDAVI 352
 DB 298 WDNLLQLTYPNGASVDPDTPVAPLGTDFSGKLFGLVLSQRDHDN-----ACRSHDAVI 356
 QY 353 ATNSAKFTPKLGAIGTWEEDVHINQTKTPVGLP--ENEGFNQWTLNYSALTN 410
 DB 357 STTSKFTPKIGISGLHSITE-HVHPNQOSRFTVGVAVDENTPPQWVLPYAGSLAN 415
 QY 411 MGLAPVAPTPGEOILFFRSHIPLKGGV--ADPVIDCLLPQEWIOLYQESAPSDVA 468
 DB 416 TMLAFAVAPTPGEOILFFRSHIPLKGGV--ADPVIDCLLPQEWIOLYQESAPSDVA 475
 QY 469 LIRFTNPDTRGLRFEAKLHRSYITVANTGSRPIVVPANGYFRFTWVQNFYSLAPMGTG 528
 DB 476 LIRVYNDTGRGLRFEAKLHRSYITVANTGSRPIVVPANGYFRFTWVQNFYSLAPMGTG 535
 QY 529 NGRRRVQ 535
 DB 536 NGRRRIQ 542

RESULT 7
 ID AAB49704 standard; protein; 539 AA.

XX AAB49704;

DT 04-APR-2001 (first entry)

DE Small round structured virus protein SEQ ID 5.

XX Small round structured virus; SRSV; food poisoning.

OS Small round structured virus.

XX WO200079280-A1.

XX 28-DEC-2000.

XX 22-JUN-2000; 2000WO-JP004095.

XX

PR 22-JUN-1999; 99JP-00175928.

XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.

PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

XX WPI; 2001-080848/09.

DR N-PSDB; AAF29145.

XX Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.

PS Claim 1; Page 50-52; 84pp; Japanese.

XX This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX

SQ Sequence 539 AA;

Query Match 67.5%; Score 1919; DB 4; Length 539;
 Best Local Similarity 65.4%; Pred. No. 2.6e-167;
 Matches 353; Conservative 80; Mismatches 97; Indels 10; Gaps 5;

QY 1 MKMASNDAPSDNDGAGLVPEANNTMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
 DB 1 MKMASNDAPSDNDGAGLVPEANNTMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
 QY 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLARMYNGYAGVEVQVLLAGNAFTAGKLV 120
 DB 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLARMYNGYAGVEVQVLLAGNAFTAGKLV 120
 QY 121 FAAPVPHPLENISPGQITMFPVHVIDVRLTLEPVLLPDVNRNFFHYNQNEPRMLVA 180
 DB 121 FAAPVPHPLENISPGQITMFPVHVIDVRLTLEPVLLPDVNRNFFHYNQNEPRMLVA 180
 QY 181 MLYTPLRNSGDDVFTVSCRVLTSPDPPFNLYPPPTLESKTKPFTLPILTIGELTNS 240
 DB 181 MLYTPLRNSGDDVFTVSCRVLTSPDPPFNLYPPPTLESKTKPFTLPILTIGELTNS 240
 QY 241 RFPVPIDELTSPNESLVVQPNQRCALDGELOQTQLPTAICFRGRINQKVSNGHV 300
 DB 241 RFPVPIDELTSPNESLVVQPNQRCALDGELOQTQLPTAICFRGRINQKVSNGHV 300
 QY 301 WNMQVTNINGTPTDPTGDPVAPLGTDFSGKLFGLVLSQRD-HDNACRSHDAVIATNSAKF 359
 DB 299 YTMNLASQNSNYDPTTEIPAPLGTDFSGKLFGLVLSQRD-HDNACRSHDAVIATNSAKF 358
 QY 360 TPKLGAIGTWEEDVHINQTKTPVGLPENGFN-----QWTLNYSALTNMGL 413
 DB 359 TPKLGSVQYTTDNTNDFQTGQTKFTPVGVVQ-DGNHQNQNEPQWVLPYSLGRTGHVHL 417
 QY 414 APPVAPTPGEOILFFRSHIPLKGGVADPVIDCLLPQEWIOLYQESAPSDVALIRFT 473
 DB 418 APAVAPTPGEOILFFRSHIPLKGGVADPVIDCLLPQEWIOLYQESAPSDVALIRFT 477
 QY 474 NPDTRGLRFEAKLHRSYITVANTGSRPIVVPANGYFRFTWVQNFYSLAPMGTGNGRRR 533
 DB 478 NPDTRGLRFEAKLHRSYITVANTGSRPIVVPANGYFRFTWVQNFYSLAPMGTGNGRRR 537

RESULT 8
 AAB49710

ID AAB49710 standard; protein; 541 AA.

XX AAB49710;

XX

DT	04-APR-2001 (first entry)	104	CHDMVGTSTTTWQWRCRSGNCCSSGHRYPVPMNRVTWIV-LSHKSQGFSTSTR	464
DE	Small round structured virus protein SEQ ID 11.	105	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
KW	Small round structured virus; SRSV; food poisoning.	106	KLPLNLRLWPLIRFNPDTGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	524
XX	Small round structured virus.	107	FYSLAPMGTCGRRRVQ	535
OS	WO200079280-A1.	108	FYSLAPMGTCGRRRVQ	541
XX	28-DEC-2000.	109	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
XX	22-JUN-2000; 2000WO-JP004095.	110	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
XX	22-JUN-1999; 99JP-00175928.	111	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
XX	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.	112	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
PA	(DENK-) DENKA SEIKEN KK.	113	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
XX	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;	114	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
PI	WPI; 2001-080848/09.	115	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
DR	N-PSDB; AAF29151.	116	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
DR	Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.	117	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
XX	Claim 1; Page 64-66; 84pp; Japanese.	118	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
XX	This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks	119	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
XX	Sequence 541 AA;	120	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
XX	Query Match 58.2%; Score 1656; DB 4; Length 541;	121	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
XX	Best Local Similarity 59.2%; Pred. No. 4.2e-143;	122	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
XX	Matches 330; Conservative 67; Mismatches 122; Indels 38; Gaps 7;	123	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
QY	1 MKMASNDAAAPSDGAGLVPEANNETWALBPVAGASTAAPTGGNNIIDPWIRNFVQAP	60	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
DB	1 MKMASNDAAAPSDGAGLVPEANNETWALBPVAGASTAAPTGGNNIIDPWIRNFVQAP	60	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
QY	61 NGEFTVSPRNSPGEVLLNLELGPENLPYLALHLSRMVYAGGVEVQVLLAGNAFTAGLV	120	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
DB	61 AGEFTVSPRNSPGEVLLNLELGPENLPYLALHLSRMVYAGGVEVQVLLAGNAFTAGLV	120	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
QY	121 FAAVPPHPPLENISPQGITMFPPIVVDVRLPVLPLDVRNFFHYNQNEPRMLVA	180	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
DB	121 FAAPPPGPFYENLSPSQITMCPHVIDVRQLEPFLPMDIWNFFHYNQNEPRMLVA	180	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
QY	181 MLYPLRNGSGDDVFTVSCRVLPDPFNVLVPTLESKTKPTLITLIGELTNS	240	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
DB	181 MLYPLRNGSGDDVFTVSCRVLPDPFNVLVPTLESKTKPTLITLIGELTNS	240	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
QY	241 RFPVDELYTSPNLSVQVQNGRCALDGLQGTQLLPTAICSPFRINQKVSQE--N	298	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
DB	241 RFPVDELYTSPNLSVQVQNGRCALDGLQGTQLLPTAICSPFRINQKVSQE--N	298	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
QY	299 HVMNQVNTNNGTTPDPTGDPAPLGTDFSGKLPGLVLSQR-----DHDNACHSHDAVTN	355	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
DB	300 YRMDEMTNTDCTPIDPTEDTGPICSPDFQGLFEGVASQRNKNQNEQNPATRAHEAINTG	359	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
QY	356 SAKFTPKLGAIOGTWEEDDVHINQTKPTPGLFENEGNCQWTLPNYSGALT-----	409	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
DB	360 GDHLCPQISSSIYITSPNLRCTNPQPLQPSGL-----RGTLIRSDNGH	405	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518
QY	410 --NMGLAPPVAPTFEGEQLIFRSHIPLKGGVADPVIDCLLPQEWIQLHYESAPSQSD--	466	-----VALIRFTNPTDGRVLFPAKLRHSYITVANTGSPPIVVPANGYERFDTWNO	518

QY 177 RLVMALYTPLRNGSGDDVFTVSCRVLTTRSPDPDFNYLVPTLESKTKPFTLPILTIGE 236
 DB 181 RLVCMLYTPLRNGSGDDVFTVSCRVLTTRSPDPDFNYLVPTLESKTKPFTLPILTIGE 240
 QY 237 LTNSRFPVPIDELITYSPNESLVVQONGRCALDGELOQTTLPTAICSFGRINQKYS 296
 DB 241 LNSRAPLPISGMGISPONVQVQNGRCALDGELOQTTLPTAICSFGRINQKYS 298
 QY 297 ENHVMNQVNTINGPDPFTGDPVAPLGTDPFSG-----KLFGLVSQRDHDNACRSHD 349
 DB 299 -----INLTDLGTFFHFF-EGPAPIGFPDLGGCDWHNMVQFGHSSQTYD----- 344
 QY 350 AVIATNSAKFTPKLGAIQ-----IG--TWEEDDVHINQFTKFTPVGLFENGSGFNQW 398
 DB 345 --VDTTPTDFVPHLGSIQANGISGNYIGVLSWVSPSH-----PSG-----SQVDLW 390
 QY 399 TLPNYSGALTLNMGALPPVAPTFPGEQILFFRSHIPLKGVADPVIDCLLPQEWIHLVQ 458
 DB 391 KIPNYGSSITEATHLAPSVYPPGFGVILFFMFKIPGPGAYSLP---CLLPQEWIHLVQ 447
 QY 459 ESAPQSDVALIRFTNPDTRGTVLFEAKLHRSYITVANTG--SRPIVVPANGYFRFDTWV 516
 DB 448 EQAPTVEAALLHYVDPDTGRITLGEFKAYPDGFLTCVPGNGASSGQQPLPINGVVFVSVV 507
 QY 517 NOFYSLAPMGNGRR 532
 DB 508 SRFYQLKPVGTASSAR 523

RESULT 10

AAR50972
 ID AAR50972 standard; protein; 530 AA.

AC AAR50972;

DT 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 05-OCT-1994 (first entry)

XX Norwalk virus strain 8FIIa protein (encoded by ORF2).

XX Norwalk virus; pathogen; acute gastroenteritis; food poisoning;
 KW seafood contamination; diagnostic assay; calicivirus; small round virus.

XX OS Norwalk virus; (strain 8FIIa).

XX PN W09405700-A2.

PD 17-MAR-1994.

PF 07-SEP-1993; 93WO-US008447.

PR 07-SEP-1992; 92US-00941365.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX PA Matson DO, Estes MK, Jiang X, Graham DY;

XX WPI; 1994-101125/12.

DR N-PSDB; AAQ56826.

XX DNA from Norwalk and related viruses - used for preparing prods. for use
 PT in diagnostic assays, detection and vaccines for Norwalk and related
 PT viruses.

XX Claim 14; Page 68-70; 156pp; English.

XX The Norwalk virus was isolated from stool samples from adult volunteers
 CC infected with safety tested Norwalk virus strain 8FIIa. The coding
 CC sequence is useful for the design of probes for use in diagnostic assays
 CC for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct
 CC PN field.) (Updated on 16-OCT-2003 to standardise OS field)

XX SQ Sequence 530 AA;
 Query Match 42.1%; Score 1196.5; DB 2; Length 530;
 Best Local Similarity 46.4%; Pred. No. 8.1e-101;
 Matches 254; Conservative 81; Mismatches 172; Indels 41; Gaps 10;
 QY 1 MKMAANDAPNSVDGAAG---LVPEAN-NETMALBPVAGASIAAPITGQNNIIDPWIRLNF 56
 DB 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSTAVATAGQVNPIDPWINN 60
 QY 57 VOAPNGEFTVSPRNGSGEVLLNLELGPDLNLAHLSRMYNGVAGGVEVQVLLAGNAFTA 116
 DB 61 VOAPNGEFTVSPRNGSGEVLLNLELGPDLNLAHLSRMYNGVAGGVEVQVLLAGNAFTA 120
 QY 117 GKLVFAAAPPFHPPLNENISPGQITMPPHVIIDVRLTEPVLPLPDRVNNFFHYNQNEPRM 176
 DB 121 GKIIVCIPPGCGSHNLITIAQATLFPHVIAVRLDPIEVLDPVLEDRVNLFHNDRNQOTM 180
 QY 177 RLVMALYTPLRNGSGDDVFTVSCRVLTTRSPDPDFNYLVPTLESKTKPFTLPILTIGE 236
 DB 181 RLVCMLYTPLRNGSGDDVFTVSCRVLTTRSPDPDFNYLVPTLESKTKPFTLPILTIGE 240
 QY 237 LTNSRFPVPIDELITYSPNESLVVQONGRCALDGELOQTTLPTAICSFGRINQKYS 296
 DB 241 LNSRAPLPISGMGISPONVQVQNGRCALDGELOQTTLPTAICSFGRINQKYS 298
 QY 297 ENHVMNQVNTINGPDPFTGDPVAPLGTDPFSG-----KLFGLVSQRDHDNACRSHD 349
 DB 299 -----INLTDLGTFFHFF-EGPAPIGFPDLGGCDWHNMVQFGHSSQTYD----- 344
 QY 350 AVIATNSAKFTPKLGAIQ-----IGTWEEDDVHINQFTKFTPVGLFENGSGFNQW 406
 DB 345 --VDTTPTDFVPHLGSIQANGISGNYIGVLSWVSPSH-----PSG-----SQVDLW 390
 QY 407 LTNMGALPPVAPTFPGEQILFFRSHIPLKGVADPVIDCLLPQEWIHLVQ 458
 DB 399 ITEATHLAPSVYPPGFGVILFFMFKIPGPGAYSLP---CLLPQEWIHLVQ 455
 QY 467 VALIRFTNPDTRGTVLFEAKLHRSYITVANTG--SRPIVVPANGYFRFDTWV 516
 DB 456 AALLHYVDPDTGRITLGEFKAYPDGFLTCVPGNGASSGQQPLPINGVVFVSVV 515
 QY 525 MGTGNRR 532
 DB 516 VGTASSAR 523
 RESULT 11
 ADC72176
 ID ADC72176 standard; protein; 530 AA.
 AC ADC72176;
 XX 18-DEC-2003 (first entry)
 DT Norwalk virus protein 2 amino acid sequence.
 DE immune response; non-Norwalk virus agent; immunogen; Norwalk virus;
 KW viral pathogen; acute gastroenteritis; virucidal; antiviral vaccine.
 XX Norwalk virus.
 XX US6572862-B1.
 XX 03-JUN-2003.
 XX 07-JUN-1995; 95US-00486049.
 XX 08-NOV-1989; 89US-00433492.
 XX 27-APR-1990; 90US-00515993.
 XX 27-AUG-1990; 90US-00573509.
 XX 06-MAY-1991; 91US-00696454.


```

Db 352 SVQTNVQGVFVPHLGSIQDEVENHHTGDIYIGIE-----WISQPS--TPPG-----TDINLW 401
Qy 399 TLNPTNSQGLTLMNGLAPPVAPTFPGEQILFFRSHIPLKGGVADP-VIDCLLPQEWIOHLY 457
Db 402 EIPDYGSSLSQANLAPPVFPFGFGEALVYFVSAPGPNRSAPNDVPCLLPQEIYTHFV 461
Qy 458 QESAPSQSDVALIRTNPDTRGVLPFAKLRHSGYITVA--NTGSRPIVVPANGYFRFTDW 515
Db 462 SEQAPTMGDAALLHYVDPDTNRLNGEFKLYPGGYLTCVPNGVGAGPQQLPLNGVLFVSW 521
Qy 516 VNQFYSLAPMGTCN-----GRRV 534
Db 522 VSRFYQLKPVGTASTARSRLGVRI 546

RESULT 13
AAB49700
ID AAB49700 standard; protein; 545 AA.
XX
AC AAB49700;
DT 04-APR-2001 (first entry)
DE Small round structured virus protein SEQ ID 1.
XX
KW Small round structured virus; SRSV; food poisoning.
XX
OS Small round structured virus.
XX
PN WO200079280-A1.
XX
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000WO-JP004095.
XX
PR 22-JUN-1999; 99JP-00175928.
XX
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
DR WPI; 2001-080848/09.
DR N-PSDB; AAF29141.
XX
PT Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
PS Claim 1; Page 40-42; 84pp; Japanese.
XX
CC This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX
SQ Sequence 545 AA;

Query Match 41.3%; Score 1174.5; DB 4; Length 545;
Best Local Similarity 45.0%; Pred. No. 9e-99;
Matches 251; Conservative 87; Mismatches 181; Indels 39; Gaps 13;

Qy 1 MKVAGNDAPSNDGAG---LVPEANN-ETMALEPVAGASIAAPLTGQNNIIDPWIRLNF 56
Db 1 MMWASKDPTNNDGTSAGQLVPEANTAEPISMFVAGATAAATAGQVNMIDPWIMNYY 60
Qy 57 VQAPNGEFTVSPRSPGFEVLLNLELGPBLNPVLAHSLRMVNGYAGGVEVQVLLAGNAFTA 116
Db 61 VQAPQGEFTISPNTTSGDILFDLQGLPHLNPFLSHLAQWYNGVGNMKYKVLLAGNAFTA 120

```

```

Qy 117 GKLVFAAVPHPPPLENISPGQITMPPHVIIDVRLTEPVLPLDVRNNPFFHNOQNEPRM 176
Db 121 KIIITSCIPPGFAAQNISIAQATMEPHVIADRVLEPIEVPLEDVRNVLFH-NNDNAPTM 179
Qy 177 RLVMALYTPLSNG--SGDDVFTVSCRVLTRPSDPDFENYLVPPPTLESKTKPFTLPILTI 234
Db 180 RLVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPSPVNDPL 239
Qy 235 GELTNSRFPVPIDELYTSPNESLVVQPQNGRCALDGLQGTTLQLLPTAICSPGRINQKV 294
Db 240 NTLNSRVPSLTKSMVSRDHQWVQFQNGRVTLQGLQGTTPTSASQLCKIRGSVFHAN 299
Qy 295 SGENHVMNQVTNINGTPTDGDVPAPLGTDPFGSKLFGVLSQSD-HDNAC-----RSH 348
Db 300 GCGY----NLTELDGSPYH-AFESAPAGIPPD-----LGECDWHMEASPTQFNFG 346
Qy 349 DAVIATN---SAKFTPKLGAIGTWEEDDVHINQTKF---TPVGLPENEGFNQWTLN 402
Db 347 DVIKQINVKQBSAFAPHLGTIQADGLSDVSVNTNMIAGLWVSPVSDGHRGVDPMVIPR 406
Qy 403 YSGALTLMNGLAPPVAPTFPGEQILFFRSHIPLKGGVADPVIDCLLPQEWIOHLYQESAP 462
Db 407 YGSTLTEAAQLAPPYIPPGFGEAIVFFMSDFPIAHGTNGLSVPCCTIPQEFVTHFVNEQAP 466
Qy 463 SQSDVALIRTNPDTRGVLPFAKLRHSGYITVA--NTGSRPIVVPANGYFRFTDWVWNPY 520
Db 467 TRGEAALLHYLDPDTHRLNGEFKLYPGFMTCPVNSSGTGPOTLPINGVFVPSVWSRFY 526
Qy 521 SLAPMGTCN-----GRRR 533
Db 527 QLKPVGTAGPACRLGIRR 544

RESULT 14
AAR57091
ID AAR57091 standard; protein; 530 AA.
XX
AC AAR57091;
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 05-OCT-1994 (first entry)
XX
DE Small round virus SRSV/KY/89 capsid protein.
XX
KW pathogen; acute gastroenteritis; food poisoning; seafood contamination;
KW diagnostic assay; human calicivirus; small round virus; SRSV; KY89;
KW Norwalk virus; capsid protein.
XX
OS Small round structured virus.
XX
PN WO9405700-A2.
XX
PD 17-MAR-1994.
XX
PF 07-SEP-1993; 93WO-US008447.
XX
PR 07-SEP-1992; 92US-00941365.
XX
PA (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI Matson DO, Estes MK, Jiang X, Graham DY;
XX
DR WPI; 1994-101125/12.
DR N-PSDB; AAQ56832.
XX
PT DNA from Norwalk and related viruses - used for preparing prods. for use
PT in diagnostic assays, detection and vaccines for Norwalk and related
PT viruses.
XX
PS Example 7; Fig 13a; 156pp; English.
XX

```


Db 529 PVGTAGPARGRL 540

Search completed: June 1, 2004, 13:45:58
Job time : 47.9597 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6.
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 12.7466 Seconds
(without alignments)
2166.837 Million cell updates/sec

Title: US-09-926-799-8
Perfect score: 2843
Sequence: 1 MMASNDAPSDGAGLVLP.....VNOFYSLAPMGTCGRRRVQ 535

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS-COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1196.5	42.1	530	4	US-08-486-049-3
2	291	10.2	622	4	US-09-590-020-6
3	290	10.2	623	4	US-09-590-020-4
4	286	10.1	623	4	US-09-590-020-2
5	284	10.0	668	4	US-09-617-594A-4
6	277.5	9.8	669	4	US-09-617-594A-2
7	273.5	9.6	626	4	US-09-590-020-7
8	122	4.3	980	2	US-08-473-553A-6
9	122	4.3	985	2	US-08-473-553A-2
10	121.5	4.3	2318	3	US-09-091-219-24
11	121.5	4.3	2318	4	US-09-660-541-24
12	113	4.0	2206	1	US-07-852-260-2
13	113	4.0	2206	2	US-08-461-503-2
14	113	4.0	2206	3	US-08-465-250-2
15	111	3.9	761	2	US-08-124-981A-2
16	110	3.9	716	1	US-08-396-479B-4
17	110	3.9	716	1	US-08-818-823-4
18	110	3.9	716	3	US-09-037-190-38
19	110	3.9	716	3	US-09-037-192-38
20	110	3.9	716	3	US-09-037-143-38
21	110	3.9	716	3	US-09-049-691-38
22	110	3.9	716	3	US-08-260-174-38
23	110	3.9	716	4	US-09-338-128A-38
24	110	3.9	716	4	US-09-232-346-38
25	110	3.9	716	4	US-09-037-192-38
26	109	3.8	716	5	PCT-US94-07297-37
27	109	3.8	913	4	US-09-623-624-2

ALIGNMENTS

RESULT 1
US-08-486-049-3
; Sequence 3, Application US/08486049
; Patent No. 6572862
; GENERAL INFORMATION:
; APPLICANT: Bates, Mary K
; APPLICANT: Jiang, Xi
; APPLICANT: Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Ave., N.W.
; CITY: Washington, D.C.
; STATE:
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,049
; FILING DATE: June 7, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311.023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-486-049-3

Query Match 42.1%; Score 1196.5; DB 4; Length 530;
Best Local Similarity 46.4%; Pred. No. 2.1e-108;
Matches 254; Conservative 81; Mismatches 172; Indels 41; Gaps 10;
QY 1 MMASNDAPSDGAG---LVPEAN-NETMALEPVAGASIAAPLTGQNIIDPWIRLNF 56
DB 1 MMASKDATSSVDGASGAGQLVPEYNASDPLAMDVPAGSVATAGVNPIDPWIRNF 60

128 PLENISPGQITMSPHVIIDVRLLEPVLPLDVRNNFFHYNQOQNEPRMRLVAMLYTPLR 187
194 -GIEPVQSTSMQPHVLFDAQVEPVLFAIPDLRLSNLYHLMSTD-TTSLVIMVYNDL- 250
188 SNGSGDDVFTVSC--RVLTRESPPDFDNVLPPTLESKTKPFTLPILTIGELTNSRPP-- 243
251 INPYANDINSSGCIIVTVETKPGDFKFHLLKPPG-----SMLTHGVSFSDLPKS 300
244 --VPIDELTYSNLSVQP--QNGRCALDGELOQTQLLPTAICSPGRINQKVS GEN 298
301 SSLWIGNRHMSDITFIIRPFVQNR-----HPDFNQETAG-- 337
299 HYNMQVNTINGTPDTPGDPAPLGTDFGKLVGSLQORDHONACRSHADAVATNSAK 358
338 --WS-----TFRFR-----PITTVSES 353
359 FTFPKGAIQIGT-----WEEDDVHINOPTKFTPVGLFE---NEGFNQWTLPNYSG 405
354 NMSKLG-IGVATDIYVFGIPDGMPTTI-----PEQITPAGIYSITASNGDITTAAGYDA 408
406 ALTL-----NMGAPPVAPTFPGEQI--LFFRSHIPLKGVADP--VID---CLLPQE 451
409 AETIVNTNPKSMYICSLQRAWDKKSINTAFITTAVRKGNISBPSNTIDMTKLWVYOD 468
452 WIOHLYQESAPSDVALIRFT-----NPDTRVLFEAKLHRSYITVANTGSRPI 502
469 --AHVGEVQTSITLALLGYTGIGBEATGSDORDKVVRSVLGETG---ARGNHPI 520

RESULT 3
US-09-590-020-4
; Sequence 4, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; PRIOR FILING DATE: 2000-06-08
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-4

Query Match 10.2%; Score 290; DB 4; Length 623;
Best Local Similarity 23.4%; Pred. No. 2e-19;
Matches 126; Conservative 79; Mismatches 189; Indels 144; Gaps 25;

13 DGAAGLVPEANNETMA----LEPVAGASTAAPLTQNNIIDPW-IRLNFVQAPNGEFTVS 67
79 EGGDGSITTPTEQGTWVGVIAPESQMSAAADMATGKSVDSWEAFFSHTSNVW----S 134
68 PRNSPGEVLLNLELGPENLYLAHLRMYNGVAGVEVQVLLAGNAFTAGKLVFAAVPPH 127
135 TSETQKILFKQSLGKPLNLYSLAKLYAVAGSVVEFRFSISGSGVFGGKLAIVVPP- 193
128 FPLENISPQITMSPHVIIDVRLLEPVLPLDVRNNFFHYNQOQNEPRMRLVAMLYTPLR 187
194 -GIEPVQSTSMQPHVLFDAQVEPVLFAIPDLRLSNLYHLMSTD-TTSLVIMVYNDL- 250
188 SNGSGDDVFTVSC--RVLTRESPPDFDNVLPPTLESKTKPFTLPILTIGELTNSRPP-- 243
251 INPYANDINSSGCIIVTVETKPGDFKFHLLKPPG-----SMLTHGVSFSDLPKS 300
244 --VPIDELTYSNLSVQP--QNGRCALDGELOQTQLLPTAICSPGRINQKVS GEN 298

57 VOAPNGEFTVSPRNSPGVLLNLELGPENLYLAHLRMYNGVAGVEVQVLLAGNAFTA 116
61 VOAPNGEFTVSPRNSPGVLLNLELGPENLYLAHLRMYNGVAGVEVQVLLAGNAFTA 120
117 GKLFAAVAPPHPLENIISPGQITMSPHVIIDVRLLEPVLPLDVRNNFFHYNQOQNEPRM 176
121 GKIIIVSCIPPGSGHNLTAQATLPHVIAVRLDPIEVLDPVRLVFNHNRNQOTM 180
177 RUVAMLYTPLRNNGSGDDVFTVSCRVLTRESPPDFDNVLPPTLESKTKPFTLPILIGE 236
181 RLVCMLYTPLRNNGSGDDVFTVSCRVLTRESPPDFDNVLPPTLESKTKPFTLPILIGE 240
237 LNSRPPVPIDELITSPNLSVQVQNGRCALDGELOQTQLLPTAICSPGRINQKVS GEN 296
241 LNSRAPLPISSMGISPDNVQVQNGRCALDGELOQTQLLPTAICSPGRINQKVS GEN 298
297 ENHVMQVNTINGTPDTPGDPAPLGTDFGKLVGSLQORDHONACRSHADAVATNSAK 358
299 --WS-----TFRFR-----PITTVSES 353
350 AVIATNSAKTPKLGAIQ--IGTWEEDDVHINOPTKFTPVGLFEENEGFNQWTLPNYSGA 406
345 --VDITPDTFVPHLGSIQANGISGN---YGVLSWISPPSHPSGQVLDWKIPNYGSS 398
407 LTLNMGAPPVAPTFPGEQIIFRSHIPLKGVADPVIDCLLPQEWIOHLYQESAPSD 466
399 ITEATHLAPSVYPPGGEVLVFFMSKMPGAYNLP---CLLPQEWIOHLYQESAPSD 455
467 VALIRFTNPDTRVLFEAKLHRSYITVANTG--SRPIVVPANGYFREDTWNQVSLAP 524
456 AALLHYVDPDTRNLGFEKAYPDGFLTCVPNGASSGQQLPFGVVFVWSRFRYQLKP 515

525 MGTGNGR 532
516 VGTASSAR 523

US-09-590-020-6
; Sequence 6, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; PRIOR FILING DATE: 2000-06-08
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-6

Query Match 10.2%; Score 291; DB 4; Length 622;
Best Local Similarity 23.6%; Pred. No. 1.6e-19;
Matches 127; Conservative 77; Mismatches 190; Indels 144; Gaps 25;

13 DGAAGLVPEANNETMA----LEPVAGASTAAPLTQNNIIDPW-IRLNFVQAPNGEFTVS 67
79 EGGDGSITTPTEQGTWVGVIAPESQMSAAADMATGKSVDSWEAFFSHTSNVW----S 134
68 PRNSPGEVLLNLELGPENLYLAHLRMYNGVAGVEVQVLLAGNAFTAGKLVFAAVPPH 127
135 TSETQKILFKQSLGKPLNLYSLAKLYAVAGSVVEFRFSISGSGVFGGKLAIVVPP- 193

APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V
; FILE REFERENCE: 454313-3151.1
; CURRENT APPLICATION NUMBER: US/09/617,594A
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-617-594A-2

Query Match 9.8%; Score 277.5; DB 4; Length 669;
Best Local Similarity 28.4%; Pred. No. 3.7e-18;
Matches 103; Conservative 49; Mismatches 152; Indels 59; Gaps 16;

QY 30 EPVAGASIAAPLTGQNNIIDPW-IRNFVQAPNGEFTVSPRNSPGEVLLNLELGPENPY 88
DB 146 EPSAQMATADAATGKSDSEWEAFSFTSVNM-----STSETQKILFKQSLGPLLNPY 201
QY 89 LAHLSRMYGAGGVEVQVLLAGNAFTAGLVFAAVPPHPFPLENISPGQIT---MPPHVI 145
DB 202 LEHLSKLYVAVSGSDVRFSGVFGKLAIVVPP-----GVDPPQSTMLQYPHVL 256
QY 146 IDVRLTEPVLPLPDVRRNFFHYNQNEPRMLVAMLYTPL-----RNSGDDVFTVS 199
DB 257 FDARQVEPVIFCLPDLKSTLYHLMSDTD-TTSLVIMVYNDLINPYANDNSGCIY---- 311
QY 200 CRVLTRSPDPDFNLYVPPTLESKTKPFTLPILTIGELTNSRPP-----VPIDELYSPE 255
DB 312 -TVETKGPDPFKHLLKPPG-----SMLTHGSPVPSDLIP 298
QY 256 SLVVOP---QNGRCALDGEIQG--TTQLLPATICSFRGRINQKVGSEHVMNQVTNI- 308
DB 361 DFVIRPFVQNRHFDNQETAGWSTFRFPITIT--SSNGSKLGTGVATDIYVGP 418
QY 309 NGTFFDPTGDPAPLGTPTDPSGKLVLSQDHD-----NACRSHDAVIATNSAKFTKLG 364
DB 419 DGWPDPTTIGELTLAG--DYS-----ITNGSGNDIATANAYDSADVITNTNFRGMVIG 471
QY 365 AIQ 367
DB 472 ALQ 474

RESULT 7
US-09-590-020-7
; Sequence 7, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 7
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Feline calicivirus

US-09-590-020-7

Query Match 9.6%; Score 273.5; DB 4; Length 626;
Best Local Similarity 32.3%; Pred. No. 8.3e-18;
Matches 71; Conservative 36; Mismatches 86; Indels 27; Gaps 8;

QY 30 EPVAGASIAAPLTGQNNIIDPW-IRNFVQAPNGEFTVSPRNSPGEVLLNLELGPENPY 88
DB 100 EPSAQMATADAATGKSDSEWEAFSFTSVNM-----STSETQKILFKQSLGPLLNPY 155
QY 89 LAHLSRMYGAGGVEVQVLLAGNAFTAGLVFAAVPPHPFPLENISPGQIT---MPPHVI 145
DB 156 LSHLAKLYVAVSGSIEVRFSGVFGKLAIVVPP-----GVDPPQSTMLQYPHVL 210
QY 146 IDVRLTEPVLPLPDVRRNFFHYNQNEPRMLVAMLYTPLRNSGDDVFTVSC--RYL 203
DB 211 FDARQVEPVIFCLPDLKSTLYHLMSDTD-TTSLVIMVYNDLINPYANDNTSSGCIYVTE 268
QY 204 TRPSDPDFNLYVPPTLESKTKPFTLPILTIGELTNSRPP 243
DB 269 TKGPDPFKHLLKPPG-----SMLTHGSPVPSDLIP 298

RESULT 8
US-08-473-553A-6
; Sequence 6, Application US/08473553A
; Patent No. 5859338
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliot M.
; APPLICANT: Clark, Steven E.
; APPLICANT: Williams, Robert W.
; TITLE OF INVENTION: Plant Clavatal Nucleic Acids,
; TITLE OF INVENTION: Transformed Plants, and Proteins
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Teet, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,553A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 980 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-473-553A-6

Query Match 4.3%; Score 122; DB 2; Length 980;
Best Local Similarity 20.9%; Pred. No. 0.012;
Matches 119; Conservative 64; Mismatches 157; Indels 230; Gaps 30;

QY 90 AHLSRM-----YNGYAGV-----EVQVLLAGNAFTAGLVFAA----- 123

Db 210 AFLSLKLNREMYIGYNSYTCGVPRFGLTKLBILDWASCTLTGTEIPTSLNKLHHT 269
QY 124 -----VPPHPLE-----NISPGQIT-MFPVHVIDVRLPEVLLPLPVRNPF 165
Db 270 LFLHNNLTGHIPELGLVSLKSLDLSINQLTGEBIPQSFNGLNITLNL-----FRNKL 325
QY 166 FHYNQNE-----PRMLVAM-----LYTPRSNGSGDDV-FTVSCRVLTRPSD----- 209
Db 326 --YGOIPEAIGELPKLEVEVWENNFTLQLPANLGRNGNLKLDVSDNHLTGLIPKDLR 383
QY 210 -----PDFNYLVPPTLE-----SKTK-----ESLVVQPNQRCALDGL----- 272
Db 384 GEKLEMLILSNFPFGPIPEELGCKSLTKIRIVKNLLNGTVPAGLFLNPLVTIITLTDN 443
QY 241 RF-----PVP-----IDELTSPN-----ESLVVQPNQRCALDGL----- 272
Db 444 FPGELPVTMSGDVLDQIYLSNNWFSGEIPPAIGNFNLQTLFLDRNFRGNIPREIPEL 503
QY 273 -----QOTTQLLPTAI--CS-----FRGRINQKV-SGENHVNMMQVTHNGTP 312
Db 504 KHLSRINTSANNITGGIPDSISRCSTLISVDLSRNRINGEIPKGINNVKNLGLTINISNQ 563
QY 313 FPGTGDVPAPLGT-----PDFSGK-----LFGVLSQRDHDNACRSHDAVIA 353
Db 564 L--TGSIFTGIGNMTSLTLDLSFNDLSGRVPLGQFLVFNETSFAGNTYLCPLHRVSCP 621
QY 354 T-----NSAKFTPK-----LGAIQIGTWEEDDVHINQPTKFTPVGLFENEG 394
Db 622 TRPGQTSDDHNTALFSPSRIVITVIAAITGLILI-----SVAIRQMNK-----KKQK 669
QY 395 FNQWTLPNYSGALTLMGLAPPVAPTFPGEQILFPRSHIPLKGGVADPVIDCLLPQEWIQ 454
Db 670 SLAWKLTAFA-----QKLDKFS-----EDVLECKEENIIG 699
QY 455 H-----LYQESAPQSODVALIRFTNPDTRGR 479
Db 700 KGGGIVYRSGMPNNVDVAIKRLVGRGTGR 729

RESULT 9
US-08-473-553A-2
; Sequence 2, Application US/08473553A
; Patent No. 5859338
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliot M.
; APPLICANT: Clark, Steven E.
; APPLICANT: Williams, Robert W.
; TITLE OF INVENTION: Plant Clavatal Nucleic Acids,
; ADDRESSSEE: Flehr, Hobbach, Test, Albritton & Herbert
; CORRESPONDENCE ADDRESS:
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,553A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-1249

; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 985 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-473-553A-2

Query Match 4.3%; Score 122; DB 2; Length 985;
Best Local Similarity 20.9%; Pred. No. 0.012;
Matches 119; Conservative 64; Mismatches 157; Indels 230; Gaps 30;

QY 90 AHLSSM-----YNGYAGV-----EVOVLLAGNAFTAGKLVFAA----- 123
Db 215 AFLSLKLNREMYIGYNSYTCGVPRFGLTKLBILDWASCTLTGTEIPTSLNKLHHT 274
QY 124 -----VPPHPLE-----NISPGQIT-MFPVHVIDVRLPEVLLPLPVRNPF 165
Db 275 LFLHNNLTGHIPELGLVSLKSLDLSINQLTGEBIPQSFNGLNITLNL-----FRNKL 330
QY 166 FHYNQNE-----PRMLVAM-----LYTPRSNGSGDDV-FTVSCRVLTRPSD----- 209
Db 331 --YGOIPEAIGELPKLEVEVWENNFTLQLPANLGRNGNLKLDVSDNHLTGLIPKDLR 388
QY 210 -----PDFNYLVPPTLE-----SKTK-----PFTLPILTIGELTNS 240
Db 389 GEKLEMLILSNFPFGPIPEELGCKSLTKIRIVKNLLNGTVPAGLFLNPLVTIITLTDN 448
QY 241 RF-----PVP-----IDELTSPN-----ESLVVQPNQRCALDGL----- 272
Db 449 FPGELPVTMSGDVLDQIYLSNNWFSGEIPPAIGNFNLQTLFLDRNFRGNIPREIPEL 508
QY 273 -----QOTTQLLPTAI--CS-----FRGRINQKV-SGENHVNMMQVTHNGTP 312
Db 509 KHLSRINTSANNITGGIPDSISRCSTLISVDLSRNRINGEIPKGINNVKNLGLTINISNQ 568
QY 313 FPGTGDVPAPLGT-----PDFSGK-----LFGVLSQRDHDNACRSHDAVIA 353
Db 569 L--TGSIFTGIGNMTSLTLDLSFNDLSGRVPLGQFLVFNETSFAGNTYLCPLHRVSCP 626
QY 354 T-----NSAKFTPK-----LGAIQIGTWEEDDVHINQPTKFTPVGLFENEG 394
Db 627 TRPGQTSDDHNTALFSPSRIVITVIAAITGLILI-----SVAIRQMNK-----KKQK 674
QY 395 FNQWTLPNYSGALTLMGLAPPVAPTFPGEQILFPRSHIPLKGGVADPVIDCLLPQEWIQ 454
Db 675 SLAWKLTAFA-----QKLDKFS-----EDVLECKEENIIG 704
QY 455 H-----LYQESAPQSODVALIRFTNPDTRGR 479
Db 705 KGGGIVYRSGMPNNVDVAIKRLVGRGTGR 734

RESULT 10
US-09-091-219-24
; Sequence 24, Application US/09091219
; Patent No. 6171592
; GENERAL INFORMATION:
; APPLICANT: STUDDERT, Michael J.
; APPLICANT: CRABB, Brendan S.
; APPLICANT: FENG, Li
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
; FILE REFERENCE: 040268/0151
; CURRENT APPLICATION NUMBER: US/09/091,219
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: PCT/AU96/00815
; EARLIER FILING DATE: 1996-12-18
; EARLIER APPLICATION NUMBER: AU PN7201
; EARLIER FILING DATE: 1995-12-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 24

RESULT 12
US-07-852-
; Sequence
; Patent N
; GENERAL
; APPL
; APPL
; APPL
; TITL
; TITL

```
RESULT 13
US-08-461-503-2
; Sequence 2, Application US/08461503
; Patent No. 5834302
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES
; TITLE OF INVENTION: FROM CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,503
; FILING DATE: 5-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-D-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-503-2

Query Match 4.0%; Score 113; DB 2; Length 2206;
Best Local Similarity 20.0%; Pred. No. 0.36;
Matches 98; Conservative 54; Mismatches 175; Indels 164; Gaps 22;

QY 48 IDPWRLNFPVQAPNG-----EFTVSPRNSPGEVLLNLELGPENLPYLAH-----LSRMYN 97
DB 389 IDTWIPLNLESTKRNMTDMYRVTLSDSADLSQPILCLSLSPAFDPRLSHTMLGEVLNYVT 448
QY 98 GYAGGEVQVLLAGNAFTAGKLVFAVPHPHLENISPCQITMFPHVIVDVTLEPVLIP 157
DB 449 HWAGSLKFTFLFCGSMATGKILVAYAPPG-AQPPTSKEAMLGTHVIMDLGLQSSCTMV 507
QY 158 LPDVNRNPFHYNQNE-----PRMLVAMLYTPLRSNGSGDDVFTVSCRVLTRPS 207
DB 508 VPWISNVYRQTQDSFTGGYISMFYQTRIVVPLSTPKMSMLG---FVSACN-----558
QY 208 PDFDFNYLVPPTLESKTKPFTLPILTIGELNRSFPVPIDELYTPNESLVVQPNQGRCA 267
DB 559 -DFSVRLLRDTT-----HISQALPQGIEDL-----TSEVAQG-----A 591
QY 268 LDGELQGTQLLPTAICSPRGINKVSGENHVMNQVNTNGTFDPTGCVPA-----321
DB 592 LTLSLPKQDLSLF-----DTKASGPAH-----SKEVPALTAVET 625
QY 322 ----PLGTPDFSGKLVLSQDHD-----NACRSHDAVIATNSAKFTPKLGAIQ 367
DB 626 GATNPL-APSDVITQTRHVQRSRSESTIESFAGAC---VAIIEDVNEQPTTQAQK- 680
QY 368 ICTWE---BDDVHINOPTKFTVPGLFENEGFNQWTLNPSGALTLNMGLA-----414

RESULT 14
US-08-465-250-2
; Sequence 2, Application US/08465250
; Patent No. 6136570
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release 1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,250
; FILING DATE: 6-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-E-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-465-250-2

Query Match 4.0%; Score 113; DB 3; Length 2206;
Best Local Similarity 20.0%; Pred. No. 0.36;
Matches 98; Conservative 54; Mismatches 175; Indels 164; Gaps 22;

QY 48 IDPWRLNFPVQAPNG-----EFTVSPRNSPGEVLLNLELGPENLPYLAH-----LSRMYN 97
DB 389 IDTWIPLNLESTKRNMTDMYRVTLSDSADLSQPILCLSLSPAFDPRLSHTMLGEVLNYVT 448
QY 98 GYAGGEVQVLLAGNAFTAGKLVFAVPHPHLENISPCQITMFPHVIVDVTLEPVLIP 157
DB 449 HWAGSLKFTFLFCGSMATGKILVAYAPPG-AQPPTSKEAMLGTHVIMDLGLQSSCTMV 507
QY 158 LPDVNRNPFHYNQNE-----PRMLVAMLYTPLRSNGSGDDVFTVSCRVLTRPS 207
DB 508 VPWISNVYRQTQDSFTGGYISMFYQTRIVVPLSTPKMSMLG---FVSACN-----558
QY 208 PDFDFNYLVPPTLESKTKPFTLPILTIGELNRSFPVPIDELYTPNESLVVQPNQGRCA 267
```

559 -DFSRLRLDTT-----HISQSALPOGIEDL-----TSEVAQG-----A 591
268 LDGELOQTQLPRAICFRGRINQKUSGENHVMNQVTNNGFPDPTGDVPA----- 321
592 LTLSPKQDLSL-----DTKASGPAH-----SKEVPALTAVET 625
322 ----PLGTPDFSGKLFVLSQRDHD-----NACRSHDAVIATNSAKFTPKLGAIQ 367
626 GATNPL-APSDTVQRHVQRSRSESTIESFFARGAC---VAIEVDNEQPTTTRAOKL- 680
368 IGTWE---EDDVHINOPTKFTPVGLFENEGNQWTLPHYSGALTNMCLA----- 414
681 FAWWRITYKDTVQLRKLEFFTYSRDME-FTFVVTAFTNA---NNGHALNQVQIMYI 736
415 PPVAPTFPEQILFFRSHPLKGGVADPVIDCLLPQEWIOHLYQES-----APS 463
737 PFGAPT-----PKSWDDYTWQTSNPSIFTYVGAAPA 768
464 QSDVALIRFTN 474
769 RISVPVUGLAN 779

RESULT 15
US-08-124-981A-2
; Sequence 2, Application US/08124981A
; Patent No. 5837840
; GENERAL INFORMATION:
; APPLICANT: Crabtree, Gerald R.
; APPLICANT: No. 5837840thorp, Jeffrey P.
; TITLE OF INVENTION: NP-AT POLYPEPTIDES AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/124,981A
; FILING DATE: 20-SEP-1993
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 5490A-226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 761 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-124-981A-2
Query Match 3.9%; Score 111; DB 2; Length 761;
Best Local Similarity 21.5%; Pred. No. 0.097;
Matches 130; Conservative 66; Mismatches 235; Indels 174; Gaps 32;
14 GAAGLVPEA--NNETMALEPVAGASTAAPTQGNIIIDFWIRLNFVQAPNGEFTVSPRNS 71
12 GARAVFETPRRQGAATSPVPPP--AAPLPAA-----AARMPESTSPFV-PSKF 57
72 PGEVLNLELGPENLPYLAHLSRMVYAGGVQVLLAGNAFTAGKLVFAAVPPH--FP 129

130 LENTSPGOITWEPHVII-----DVRTLEPVLLPLPDVRRNF-----PHYN 169
94 SNNVSPALPLPTAHSTLPAPCHNLQSTPGIIPADHPSGYGAALDGGPAGYFLSSGHTR 153
170 QQ-----NEPRMLVAMLYTFLRNGSGDDVFTVSCVLTFRPSDFDFNVLYVPRTLESKT 224
154 PDGAPALSPRIEITSCGLYHNNNQFFHDVEVEDVLPSSKRSP--STATLSLSLEAYR 211
225 KPFTL-PILTIGELT-----NSRFPVPIDELYTSNESLVVQPQN-----G 264
212 DPSCLSPASSLSRSCNSEASYESINYPIASQTSFQSPQPCVSPKTTDPEEGFPRLG 271
265 RCALDGELOQTQLPPTAICFRGRINQKUSGENHVMNQVTNNG--TPFDP-TGDVPA 321
272 ACTLLGSPQHSPTSPRASVTEESWLGARSSRPASPCNKRKYSLNGRQPPYSPHHSPTPS 331
322 PLGTPDFS---GKLFQVLSQRDHDNACRSHDAVIATNSAKFTPKLGAIQIGTWEEDDVHI 378
332 PHGSPRVSVTDDSWLGNITQ-----YTSSAIVAAINALITD--SSLDLG-----DGVPV 378
379 ----NOP---TKFTPVGLFENEGNQWTLPHYSGALTNMGLAPPVAPTFPEQIL 427
379 KSRKTLEQPPSVALKVEPVG-----EDLGSPPPPPADFAPEDYSS 418
428 FFRSHIPLKGGVADPVIDCLLPQEWIOHLYQESAPS-----QSDVALIRFTN 474
419 F--QHI-RKGGFCQYVL--AVP---QHPYQWAKPKPLSPTSYMSPRTLPAIDWQLPSHSG 469
475 PDTGRVLFEAKL-HRSGYIT-----VANTGSRPIVVPANGYFRFDTWVNFYSLAPMG 526
470 PYELRIEVQPKSHHRAHYETEGSRGAVKASAGGHPV-VQLHGYLENEPLMLQLF----IG 524
527 TGNGR 531
525 TADDR 529

Search completed: June 1, 2004, 13:58:04
Job time : 14.7466 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 34.2005 Seconds
(without alignments)
4368.312 Million cell updates/sec

Title: US-09-926-799-8
Perfect score: 2843
Sequence: 1 MKMASNDAPSDGAGLVP.....VNQFYSLAPMTGNGRRRVQ 535

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1196.5	42.1	530	14	US-10-314-739-3
2	284	10.0	668	14	Sequence 3, Appli
3	277.5	9.8	547	12	Sequence 4, Appli
4	277.5	9.8	669	14	Sequence 4, Appli
5	277.5	9.8	671	12	Sequence 2, Appli
6	125.5	4.4	1147	15	Sequence 2, Appli
7	117	4.1	338	10	Sequence 38, Appli
8	110.5	3.9	2468	12	Sequence 2, Appli
9	110.5	3.9	2468	14	Sequence 66335, A
10	110	3.9	716	10	Sequence 4, Appli
11	110	3.9	716	14	Sequence 38, Appli
12	110	3.9	716	15	Sequence 38, Appli
13	109.5	3.9	925	15	Sequence 182, App
14	109	3.8	913	14	Sequence 3889, Ap
15	109	3.8	913	15	Sequence 2, Appli
					Sequence 132, App

16	107	3.8	716	10	US-09-349-058-46
17	107	3.8	716	14	Sequence 46, Appli
18	106.5	3.7	951	14	Sequence 46, Appli
19	106.5	3.7	951	14	Sequence 172, App
20	106.5	3.7	951	16	Sequence 172, App
21	106.5	3.7	1592	12	Sequence 172, App
22	105.5	3.7	988	12	Sequence 319, App
23	104.5	3.7	5636	14	Sequence 144895, A
24	104.5	3.7	5636	15	Sequence 128, App
25	104.5	3.7	5636	15	Sequence 72, Appli
26	102.5	3.6	548	11	Sequence 93, Appli
27	102.5	3.6	548	11	Sequence 2916, Ap
28	102.5	3.6	3930	12	Sequence 48817, A
29	102	3.6	643	14	Sequence 12113, A
30	102	3.6	2572	14	Sequence 86, Appli
31	102	3.6	2673	15	Sequence 76, Appli
32	102	3.6	3645	14	Sequence 127, App
33	102	3.6	3645	15	Sequence 73, Appli
34	102	3.6	4495	15	Sequence 20, Appli
35	102	3.6	5635	16	Sequence 78, Appli
36	101.5	3.5	1382	14	Sequence 18, Appli
37	100.5	3.5	836	14	Sequence 40, Appli
38	100.5	3.5	980	14	Sequence 41, Appli
39	100.5	3.5	2227	9	Sequence 12, Appli
40	100.5	3.5	2227	13	Sequence 12, Appli
41	100.5	3.5	2227	13	Sequence 2, Appli
42	100.5	3.5	2227	13	Sequence 4, Appli
43	100.5	3.5	2227	13	Sequence 6, Appli
44	100.5	3.5	2227	16	Sequence 12, Appli
45	100	3.5	1196	16	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-10-314-739-3
; Sequence 3, Application US/10314739
; Publication No. US20030129588A1
; GENERAL INFORMATION:
; APPLICANT: Bates, Mary K
; Jiang, Xi
; Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; Characterize No. US20030129588A1walk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pubbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Ave., N.W.
; CITY: Washington, D.C.
; STATE: <Unknown>
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/10/314,739
; APPLICATION NUMBER: US/10/314,739
; FILING DATE: 09-Dec-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,049
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311.023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX: <Unknown>

```

;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-314-739-3

Query Match 42.1%; Score 1196.5; DB 14; Length 530;
Best Local Similarity 46.4%; Pred. No. 3.3e-105; Indels 41; Gaps 10;
Matches 254; Conservative 81; Mismatches 172;

QY 1 MKWASNDAPSDGAG---LVPEAN-NETMALEPVAGASIAAPLTGQNNIIDPWIRLNF 56
DB 1 MMWASKDATSSVDGAGAGQLPEVNASDPLANDPVAGSSTAVATAGVNPIDPWIIINF 60

QY 57 VOAPNGEFTVSPRNSPEVLLNLELGPENLYLAHLSRMVNGYAGGVEQVLLAGNAFTA 116
DB 61 VQAPQGEFTISPNTGDFLDSLGLPHLNPFLHLSQMYNGWVGNMVRIMLAGNAFTA 120

QY 117 GKLVFAAVPPHPLENISPGQITMFPFHIIDVRLTLEPVLPLPDVRRNFFHYNQNEPRM 176
DB 121 GKIIVCIPPGFSGHNLTTAATLFPFHVIAVRLDPIEVPLEDVRNVLFNHNDRNQTM 180

QY 177 RLVMALYTPLRNSGDDVFTVSCRVLTRPSDFDFNYLVPPPTLESKTKPFTLPILTIOE 236
DB 181 RLVMALYTPLRNSGDDVFTVSCRVLTRPSDFDFNYLVPPPTLESKTKPFTLPILTIOE 240

QY 237 LTNSRFPVPEIDELTSPNESLVVQPNQRCALDGELOQTLLPTAICSFGRINQKVS 296
DB 241 LNSRRLPISMGISPDNVQSVQFQNGRCRLDGLVGTTPVSLSHVAKIRGTSNGTV-- 298

QY 297 ENHVMNQVNTINGTPDPTGDPAPLGTDFSG-----KLFGLVSRQDHDNACRSHD 349
DB 299 -----INLTEDDGTTPHPH-EGPAPIGFDFLGGCDWHINMTQFGHSSQTQYD----- 344

QY 350 AVIATNSAKFTPKLGAIQ---IGWEEDDVHINQPTKFTPVGLFENEGFNQWTLPNYSGA 406
DB 345 --VDTPTDFVPHLGSIQANGIGSGN---YVGVLWSISPSPHSGSQVDLWKIPNYGS 398

QY 407 LTNMGLAPVAPTFPGEQILFFESHPLKGSVADPVIDCLLPQEWIQLHYQSAPSQSD 466
DB 399 ITEATHLAPSYPPGFGVLVFFWMSKMPGAYNLP---CLLPQEYISHLASEQAPTVGE 455

QY 467 VALIRFTNPDTGRVLFEAKLHRSYITVANTG--SRPIVVPANGYFRFDTWNQFYSLAP 524
DB 456 AALLHYVDPTGRNLGKFAKYPDGLTCVPNGASSGPQQLPENGVFVSVGRFYQLKP 515

QY 525 MGTGNRR 532
DB 516 VGTASSAR 523

RESULT 2
US-10-209-507-4
; Sequence 4, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
; US-10-209-507-4

Query Match 10.0%; Score 284; DB 14; Length 668;
Best Local Similarity 26.2%; Pred. No. 9.6e-18; Indels 136; Gaps 21;
Matches 110; Conservative 49; Mismatches 125;

QY 17 GLVPEANNETMALEPVAGASIAAPLTGQNNIIDPW-IRLNFVQAPNGEFTVSPRNSPCEV 75
DB 141 GVIAEFNAQMSAVADVA-----TGK-SVDSEAEAFSFTSVNM-----STSETQSKI 187

QY 76 LLNLELGPENLYLAHLSRMVNGYAGGVEQVLLAGNAFTAGKLVFAAVAPPHPLENISP 135
DB 188 LFKQSLGSLNPNLTHLAKLYLVANGSIEVRISISGSGVFGKLAIVVPP-----GIDP 242

QY 136 GQIT--MPPHVIIDVRLTLEPVLPLPDVRRNFFHYNQNEPRMLVAMLYTPL----- 186
DB 243 VQSTSMLOYPHVLFDARQVEPVITFDPLRNSLYHLMSDTP-TTSLVIMYNDLNPVAN 301

QY 187 RNSGDDVFTVSCRVLTRPSDFDFNYLVPPPTLESKTKPFTLPILTIGELTNSRFP-- 243
DB 302 DSNSSGCV-----TVETKPGDPFKHLLKPPG-----SMLTGSIFSDLIPKSS 346

QY 244 -VPIDEITYSPNESLVVQPN---QNGRCALDGELOQTLLPTAICSFGRINQKVS 299
DB 347 SLMTIGNRHNSDITDFVILKPFVQANR-----HFDNFQETAG--- 382

QY 300 VMNQVNTINGTPDPTGDPAPLGTDFSGKLVGLVSRQDHDNACRSHDAVIATNSAKF 359
DB 383 -WS-----TPRFR-----PITIVSEKG 399

QY 360 TPKLGAIQGT-----WEEDDVHINQPTKFTPVG--LFENEGFNQ-WTLPNYSGA 406
DB 400 GSKLG-IGVATDSIVPGIPDGWPDFTI---PEKLIPAGDYAITNGNNDITTAADYDGA 454

RESULT 3
US-10-670-695-4
; Sequence 4, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Feline calicivirus
; US-10-670-695-4

Query Match 9.8%; Score 277.5; DB 12; Length 547;
Best Local Similarity 31.4%; Pred. No. 2.9e-17; Indels 31; Gaps 9;
Matches 76; Conservative 38; Mismatches 97;

QY 30 EPVAGASIAAPLTGQNNIIDPW-IRLNFVQAPNGEFTVSPRNSPCEVLLNLGLPELNEY 88
DB 21 EPSAQMSAADMATGSKSVDSWEAEAFSFTSVNM-----STSETQSKIPLKQSLGSLNPNY 76
QY 89 LAHLSRMVNGYAGGVEQVLLAGNAFTAGKLVFAAVAPPHPLENISPQGIT---MPPHVI 145
```



```
Db 77 LEHLAKLYVWAGSGSEVRSISGSGVFGGKLAIVVPP-----GVDPVQSTSMQLQPHVL 131
Qy 146 IDVRTLEPVLPLPDVRNFFHYNQNEPRMELVAMLYTPLRSNGSGDDVFTVSC--RVL 203
Db 132 FDAQVEPVIFCLPDLRSTLYHMSDTD--TTSLVIMVYNDL--INPYANDANSSGGCIVTVE 189
Qy 204 TRPSDPDFNVLVPPTLSKTKPFTLPILTIGELTNSRPP-----VPIDELYTSNPSLVV 259
Db 190 TKPGDPDFKPHLLKPPG-----SMLTHGSIPLDIPKTSLSLWIGNRYMSDITDFVI 239
Qy 260 QP 261
Db 240 RP 241

RESULT 4
US-10-209-507-2
; Sequence 2, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-2

Query Match 9.8%; Score 277.5; DB 14; Length 669;
Best Local Similarity 28.4%; Pred. No. 46-17;
Matches 103; Conservative 49; Mismatches 152; Indels 59; Gaps 16;

Qy 30 EPVAGASIAAPLTGQNNIIDPW--IRLNFVQAPNGEFTVSPRNSPGEVLLNLELGPENLPY 88
Db 146 EPSAQWATAADAATGKSDSEWESFHTSYNW----STSETQKILFKQSLGPLLNPY 201
Qy 89 LAHLSRMVYAGGVEQVLLAGNAFTAGKLVFAAVPPHFPLENTSPGQIT---MPPHVI 145
Db 202 LEHLSKLYVWAGSDVDRFSISGSGVFGGKLAIVVPP-----GVDPVQSTSMQLQPHVL 256
Qy 146 IDVRTLEPVLPLPDVRNFFHYNQNEPRMELVAMLYTPL-----RSNGSGDDVFTVTS 199
Db 257 FDAQVEPVIFIPDLRSTLYHMSDTD--TTSLVIMVYNDL--INPYANDANSSGGCIV--- 311
Qy 200 CRVLTRPSDPDFNVLVPPTLSKTKPFTLPILTIGELTNSRPP-----VPIDELYTSNPE 255
Db 312 -TVETKPGDPDFKPHLLKPPG-----SMLTHGSIPLDIPKTSLSLWIGNRYMSDIT 360
Qy 256 SLVVOV---QNGRCALDCELOG--TTQLLPATICSFRGRINQKVSNGENHVNMOVTNI- 308
Db 361 DFVIRFPVQANRHDFDQETAGWTPRPRPIITII--SENGSKLGTGVATDIYVGPLP 418
Qy 309 NGTFPDPTGDVPAPLGTDFSGKLGVLGSORDH-----NACRSHDAVIATNSAKFTPKLG 364
Db 419 DGMWPTTIGEELTPAG--DYS-----ITNGSGNDIATANAYDSADVIITNTFRGMYICG 471
Qy 365 AIQ 367
Db 472 ALQ 474
```

```
RESULT 5
US-10-670-695-2
; Sequence 2, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 671
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-670-695-2

Query Match 9.8%; Score 277.5; DB 12; Length 671;
Best Local Similarity 31.4%; Pred. No. 4,1e-17;
Matches 76; Conservative 38; Mismatches 97; Indels 31; Gaps 9;

Qy 30 EPVAGASIAAPLTGQNNIIDPW--IRLNFVQAPNGEFTVSPRNSPGEVLLNLELGPENLPY 88
Db 145 EPSAQWATAADAATGKSDSEWESFHTSYNW----STSETQKILFKQSLGPLLNPY 200
Qy 89 LAHLSRMVYAGGVEQVLLAGNAFTAGKLVFAAVPPHFPLENTSPGQIT---MPPHVI 145
Db 201 LEHLAKLYVWAGSGSEVRSISGSGVFGGKLAIVVPP-----GVDPVQSTSMQLQPHVL 255
Qy 146 IDVRTLEPVLPLPDVRNFFHYNQNEPRMELVAMLYTPLRSNGSGDDVFTVSC--RVL 203
Db 256 FDAQVEPVIFCLPDLRSTLYHMSDTD--TTSLVIMVYNDL--INPYANDANSSGGCIVTVE 313
Qy 204 TRPSDPDFNVLVPPTLSKTKPFTLPILTIGELTNSRPP-----VPIDELYTSNPSLVV 259
Db 314 TKPGDPDFKPHLLKPPG-----SMLTHGSIPLDIPKTSLSLWIGNRYMSDITDFVI 363
Qy 260 QP 261
Db 364 RP 365
```

```
RESULT 6
US-10-327-481A-38
; Sequence 38, Application US/10327481A
; Publication No. US20040001864A1
; GENERAL INFORMATION:
; APPLICANT: Burman, Alison J.
; APPLICANT: Audonnet, Jean-Christophe F.
; APPLICANT: Lombard, Michel F.A.
; TITLE OF INVENTION: Vaccine against Foot-and-Mouth Disease
; FILE REFERENCE: 454313-3178
; CURRENT APPLICATION NUMBER: US/10/327,481A
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: PCT/FR01/02042
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: FR 0008437
; PRIOR FILING DATE: 2000-06-29
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 38
; LENGTH: 1147
; TYPE: PRT
```

Wed Jun 2 09:13:33 2004

ORGANISM: Foot-and-mouth disease virus
US-10-327-481A-38

Query Match 4.4%; Score 125.5; DB 15; Length 1147;
Best Local Similarity 22.1%; Pred. No. 0.033;
Matches 71; Conservative 34; Mismatches 139; Indels 77; Gaps 10;

QY 70 NSPGEVLLNLELGPENLYLAHLNRYNGYAGGVEVQVLLAGNAFTAGKLVFAVPPHFP 129
DB 158 DRPFGYLTAKLEPDTDHGFGHLDVSAYMRNGWDEVSAGVNGFNGCLLVAVPWEKA 217
QY 130 LENISPGQITMPEPHVIDRTLEPVLPLPDRVNNFFHYNQNEPRMELVAMLYTPLRSN 189
DB 218 FDTREKQITLPHQFISPTNMTAHTVPLYGVN--RYDQY----- 257
QY 190 GSGDDVFTVSCRVLTRPSDFDNVLPPTLESKTKPTLPIITIGELTNSRFPVIDEL 249
DB 258 -----KCHKPWLTVVWVLSPLTVSVNTAAPQIKV 285
QY 250 YTSNESLVVQPQNGRCALDGELOGTQTLTALCSFRG-----RINQKVSGENHVN 302
DB 286 YAN-----IAPYVHVA--GELPSKEGIFPVACADGYGGLVTTDPKTAADPVYGR--VYN 335
QY 303 MQVTNINGTPTDTPVPAPLGTPTDF-----SGKLFGLVLSQRDHDNACRSHDAVIATNSAK 358
DB 336 PKTNYFGR--FTNLLDVAE--ACPTFLRFDDGRPY-VVTRADDTRLLAKFDVSLAAGHMS 391
QY 359 FPKLGAIQGTWEEDDVHIN 379
DB 392 NTVLSGIAQYQYQYSGTINLH 412

RESULT 7

US-09-867-932A-2
; Sequence 2, Application US/09867932A
; Publication No. US20030049825A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Han-You
; APPLICANT: Kuo, Tsun-Yung
; APPLICANT: Huang, Hsiao-I
; APPLICANT: Yang, Huey-Lang
; TITLE OF INVENTION: NERVOUS NECROSIS VIRUS PROTEIN
; FILE REFERENCE: 08919-023001
; CURRENT APPLICATION NUMBER: US/09/867,932A
; CURRENT FILING DATE: 2001-05-30
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Nervous necrosis virus
US-09-867-932A-2

Query Match 4.1%; Score 117; DB 10; Length 338;
Best Local Similarity 24.5%; Pred. No. 0.031;
Matches 80; Conservative 49; Mismatches 132; Indels 66; Gaps 18;

QY 8 AAPSDGAGLVP--EANNETWALEPVGASIAAPLT-----GNNI-IDPWIRLNFVQAP 60
DB 10 AKPATTKAANPQPRRANRRNRRTDAPVSKASTVTGRTNDVHLSGMSRISQAVLP 69
QY 61 NGEFTVSPRNSPGEVLLNLGLPELNPYLAHLNRYNGYAGGV---EVOVLLAGNAFTAG 117
DB 70 AGTGT-----DGVVVVDATIVDPLLRGLHAARIFQYAVETLEPEIQMCPAN--TGG 121
QY 118 KLVAFAVPPHFPLENTISP-QQITWPHVII-----DVRTLEPVLPLPDRVNNFFHYNQON 172
DB 122 GYVAGFLPD--PTDNDHTFGALQATRGAVAKWESRTVR-----PQYTRILLWTSSCK 173
QY 173 EPRM-----RLVAMLYTLRNSGGDDVFTVS---CRVLTRESPDFDNVLPPTLESKTKP 226
DB 174 EQLTSPGRLLILCV-----GNNTDVNVSVLCRSVR-----LSVPSLETPEET 218

QY 227 FTLPIITIGELTNSRFPVIDELYTSNESLVVQPQNGRCALDGELOGTQTLTALCSF 286
DB 219 -TAPIMTQGSYLN-----DSLSTNDSKSLGSTPLDIAPDGAVFQLDRLLSIDYSLG 270
QY 287 RGRINQKVSGENHVNMQVTNIN-GTP 312
DB 271 TGDVDRAV-----YWHLKKFAGNAGTP 292

RESULT 8

US-10-282-122A-66335
; Sequence 66335, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66335
; LENGTH: 2468
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-66335

Query Match 3.9%; Score 110.5; DB 12; Length 2468;
Best Local Similarity 18.7%; Pred. No. 3.1;
Matches 132; Conservative 64; Mismatches 246; Indels 265; Gaps 30;

QY 8 AAPSDGAGLVP--EANNETWALEPVGASIAAPLTGQNNIIDPWIRLNFVQAPNGCF 64
DB 1234 AATTVDAVAPPVDPISNGTTISGTAEAGAKVI--LTDGNG-----NPIGT 1279
QY 65 TVSPRN-----SPEVLLNLGLPELNPYLAHLNRYNGYAGGVQVLL---AGNAFTAG 117
DB 1280 TADSGSNWSTPGTPLAN-----GTVYNAVQADPAGNTGPGQ 1316
QY 118 KLVAFAVPPHFPLENTISPGQI---TWPHVIDVRTL-----EPVLLPLPDRVNNFFHYN 169

Db 1317 STTDAVAPNTVVPNSGNLLNGTAEPGSTV---TLTDGNGNPICQTTADSGNWSFTP 1373
Qy 170 QONEPRMLVAMLYTPLRNSGDDVFTVSCRVLTRPSDFDFNYLVPPTLSEKTKPFTL 229
Db 1374 GSQLPNGTVVNTASDAAGNTSLPATTVDSSLSIPQVDPNSGISCTADAGN---TI 1430
Qy 230 PILT-----IGELT-----NSRF-----PVP-----IDELYTSPNE----- 255
Db 1431 -IITDNGNPIGOVTDAGSNWSFTPGIPLDGTVVNVVARSFNSVDSAPAVITVDGVAP 1489
Qy 256 -SLVQOPQNGRCALDELQGTQLLPTAICSFRG--RINQKVSNGENHWNM-----QV 305
Db 1490 AAPVIDPSNGT-----EISGTAEAGATVILTDGGNPIGOATADSGNWTFTPSTPLANG 1544
Qy 306 TNINGTPPDPPTGDV-----PAPL----- 323
Db 1545 TVINAVAQDPAGNTSGPASVTDATAPAPVINPNSGVVISGTAEAGATVILTDGNGNPI 1604
Qy 324 -----GTPDFSGKLFVLSQRDHDNACRSHDAVI-----ATNSAKFTPKLGA 365
Db 1605 GOVTADSGNWSFTPGTPLANGSVINALAQDAAGNNSPTSATVDSLAPAAPVIDPSNGS 1664
Qy 366 IQIGTWEEDDVHINQPTKTPVGLFENEGFNOWTLP-----NYSG 405
Db 1665 VIAGTAEAGATVILTDGNGNPIGOVTDAGSNWSFTPGTPLNGTVVNAVAQDAAGNTSG 1724
Qy 406 ALTNMGLAPPVAPTF-----PGSQILF-----FRSH 432
Db 1725 PVSTTVDAVAPATPVIDPSNGVELSGTAEPGVRVILTDGNGNPIGOVTLADSGNWSFTPG 1784
Qy 433 IPLKGGVA-----DPVIDCLLPQEWIQLYQESAPSDVALIR-----FTNPDTGRVLP 482
Db 1785 TPLANGTVVNAVAQDPA-----GNTSGPASTTVDVAPATPVINPNSGVI- 1830
Qy 483 EAKLHRSGITVANTGSRPIVVPANGYFRFDTWVNOFYSLAPMGTCN 529
Db 1831 -----TGTAEVGAKVILTDGNG-----NPIGETTADGSGN 1860

RESULT 9
US-10-246-330-4
; Sequence 4, Application US/10246330
; Publication No. US20030166030A1
; GENERAL INFORMATION:
; APPLICANT: O'Toole, George A.
; TITLE OF INVENTION: METHODS TO STUDY AND MECHANISMS OF
; FILE REFERENCE: 14537-002001
; CURRENT APPLICATION NUMBER: US/10/246,330
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 60/323,241
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 2468
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-246-330-4

Query Match 3.9%; Score 110.5; DB 14; Length 2468;
Best Local Similarity 18.7%; Pred. No. 3.1;
Matches 132; Conservative 64; Mismatches 246; Indels 265; Gaps 30;

Qy 8 AAPSNDGAAGLVP---EANNETWALEPVGASIAAPLTQCNIIIDPWIRLNFVQAPNGEF 64
Db 1234 AATTVDVAPPAPVIDPSNGTGTISGTAEAGAKVI--LTDGNG-----NPIGET 1279
Qy 65 TVSPRN-----SPGEVLLNLELGPENLPYLAHLRMYNGYAGGVEVQVLL---AGNAFTAG 117
Db 1280 TADSGNWSFTPGTPLAN-----GTVVNAVAQDPAGTNGPQG 1316

Qy 118 KLVFAAVPHPPPLENISPGOI---TMFPHVIVIDVTL-----EPVLLPLPDVRNPPHYN 169
Db 1317 STTDAVAPNTVVPNSGNLLNGTAEPGSTV---TLTDGNGNPICQTTADSGNWSFTP 1373
Qy 170 QONEPRMLVAMLYTPLRNSGDDVFTVSCRVLTRPSDFDFNYLVPPTLSEKTKPFTL 229
Db 1374 GSQLPNGTVVNTASDAAGNTSLPATTVDSSLSIPQVDPNSGISCTADAGN---TI 1430
Qy 230 PILT-----IGELT-----NSRF-----PVP-----IDELYTSPNE----- 255
Db 1431 -IITDNGNPIGOVTDAGSNWSFTPGIPLDGTVVNVVARSFNSVDSAPAVITVDGVAP 1489
Qy 256 -SLVQOPQNGRCALDELQGTQLLPTAICSFRG--RINQKVSNGENHWNM-----QV 305
Db 1490 AAPVIDPSNGT-----EISGTAEAGATVILTDGGNPIGOATADSGNWTFTPSTPLANG 1544
Qy 306 TNINGTPPDPPTGDV-----PAPL----- 323
Db 1545 TVINAVAQDPAGNTSGPASVTDATAPAPVINPNSGVVISGTAEAGATVILTDGNGNPI 1604
Qy 324 -----GTPDFSGKLFVLSQRDHDNACRSHDAVI-----ATNSAKFTPKLGA 365
Db 1605 GOVTADSGNWSFTPGTPLANGSVINALAQDAAGNNSPTSATVDSLAPAAPVIDPSNGS 1664
Qy 366 IQIGTWEEDDVHINQPTKTPVGLFENEGFNOWTLP-----NYSG 405
Db 1665 VIAGTAEAGATVILTDGNGNPIGOVTDAGSNWSFTPGTPLNGTVVNAVAQDAAGNTSG 1724
Qy 406 ALTNMGLAPPVAPTF-----PGSQILF-----FRSH 432
Db 1725 PVSTTVDAVAPATPVIDPSNGVELSGTAEPGVRVILTDGNGNPIGOVTLADSGNWSFTPG 1784
Qy 433 IPLKGGVA-----DPVIDCLLPQEWIQLYQESAPSDVALIR-----FTNPDTGRVLP 482
Db 1785 TPLANGTVVNAVAQDPA-----GNTSGPASTTVDVAPATPVINPNSGVI- 1830
Qy 483 EAKLHRSGITVANTGSRPIVVPANGYFRFDTWVNOFYSLAPMGTCN 529
Db 1831 -----TGTAEVGAKVILTDGNG-----NPIGETTADGSGN 1860

RESULT 10
US-09-349-058-38
; Sequence 38, Application US/09349058
; Publication No. US20030056229A1
; GENERAL INFORMATION:
; APPLICANT: Crabtree, Gerald R.
; APPLICANT: No. US20030056229A1thiop, Jeffrey P.
; TITLE OF INVENTION: NF-AT POLYPEPTIDES AND POLYNUCLEOTIDES
; FILE REFERENCE: APV-332.12
; CURRENT APPLICATION NUMBER: US/09/349,058
; CURRENT FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 08/260,174
; EARLIER FILING DATE: 1994-06-13
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 716
; TYPE: PRT
; ORGANISM: human
US-09-349-058-38

Query Match 3.9%; Score 110; DB 10; Length 716;
Best Local Similarity 21.5%; Pred. No. 0.48;
Matches 109; Conservative 62; Mismatches 201; Indels 134; Gaps 26;

Qy 111 GNAFTAGKLVFAAVPPH--FPLENISPGQITMFPFHVII-----DVRTLEPVLLPLPDVRN 163
Db 28 GPAPRAGGTMTKSAEEHGYASSNVSPALPLPTAHTSLTLPACPNLTQSTFGIIPADHPS 87
Qy 164 NF-----FHYNOQ-----NEPRMLVAMLYTPLRNSGDDVFTVSCRVL 203

Db 88 GYGAALDGCAGYFLUSSGHTPDGAPALSPRIITSCGLYHNNNOFFHDEVEDVLPS 147
Qy 204 TRPSDFDNFVLPPTLESKTKPFTL-PILTIGELT-----NSRFPVPIDELYTSPN 254
Db 148 SKRSP--STATLSLPSLEAYRDPSCLSPASSLSRSCNSEASSYESNYSPYASPTQSPW 205
Qy 255 EBLVQOPN-----GRCALDGLGTTQLLPTAICSFGRGRINQKVSNGENHVMNM 303
Db 206 QSPCVSPKTTDPEEGFPRGLGACTLLGSPQHSPTSPRASVTEESWLGARSRPASPCNK 265
Qy 304 QVTNING--TPDTP-TGDVPAPLGTDFDS--GKLFGLVLSQRDHDNACRSHDAVATNSA 357
Db 266 RYKSLNGRQPPYPSPHSPPTSPHSGPRSVTDDSLGNTTQ-----YTSSAIVAAINA 318
Qy 358 KFTPKLGAIOIGTWEEDDVIH-----NQP-----TKFTPVCLFENEGFNQWTLNYSGA 406
Db 319 LTLD--SSLDLG---DGVFVKSRTTLEQPPSVALKVEPVG----- 354
Qy 407 LTLNMGGLAPPVAPTFPEGEQILFPRSHIPKGGVADPVIDCLLPQEWI OHLYESAPS--- 463
Db 355 --EDLGSPPPPADAFEDYSSP--QHI-RKGGFCQOYL--AVP-----QHPYQWAKPKPLS 403
Qy 464 -----QSDVALIRFTNPTGRVLFPEAKL-HRSGYIT-----VANTGSRPIVVP 505
Db 404 PTSYMSPTLPALDWQLPSSHGPGYELRIEVQPKSHHRAHYETEGSRGAVKASAGGHP1-VQ 462
Qy 506 ANGYFRDFTWVNOFYSLAPMGTGNGR 531
Db 463 LHGYLENEPLMLQLF-----IGTADDR 484

RESULT 11
US-10-040-430-38
; Sequence 38, Application US/10040430
; Publication No. US20030049641A1
; GENERAL INFORMATION:
; APPLICANT: Crabtree, Gerald R.
; APPLICANT: No. US20030049641A1throp, Jeffrey P.
; APPLICANT: Ho, Steffan M.
; APPLICANT: Flanagan, William M.
; TITLE OF INVENTION: NF-AT POLYPEPTIDES AND POLYNUCLEOTIDES AND SCREENING
; FILE REFERENCE: APV-008.04
; CURRENT APPLICATION NUMBER: US/10/040.430
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: US/09/232.346
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 08/507,032
; PRIOR FILING DATE: 1995-07-31
; PRIOR APPLICATION NUMBER: 08/228,944
; PRIOR FILING DATE: 1994-04-18
; PRIOR APPLICATION NUMBER: 07/749,385
; PRIOR FILING DATE: 1991-08-22
; PRIOR APPLICATION NUMBER: 08/260,174
; PRIOR FILING DATE: 1994-06-13
; PRIOR APPLICATION NUMBER: 08/124,981
; PRIOR FILING DATE: 1993-09-20
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 716
; TYPE: PRT
; ORGANISM: human
US-10-040-430-38

Query Match 3.9%; Score 110; DB 14; Length 716;
Best Local Similarity 21.5%; Pred. No. 0.48;
Matches 109; Conservative 62; Mismatches 201; Indels 134; Gaps 26;
Qy 111 GNAFTAGKLVAFAVPPH--FPLENISPGQITMPPHVI-----DVRTLEPVLPLPDVNR 163
Db 28 GPAPRAGGTMKSAEEHYGASSNVSPALPLPTAHSTLPAPCHNLQSTPGIIPADHPS 87

Qy 164 NF-----FHYNQ-----NEPRMLVAMLYTPLRSNGSGDDVFTVSCRVL 203
Db 88 GYGAALDGCAGYFLUSSGHTPDGAPALSPRIITSCGLYHNNNOFFHDEVEDVLPS 147
Qy 204 TRPSDFDNFVLPPTLESKTKPFTL-PILTIGELT-----NSRFPVPIDELYTSPN 254
Db 148 SKRSP--STATLSLPSLEAYRDPSCLSPASSLSRSCNSEASSYESNYSPYASPTQSPW 205
Qy 255 EBLVQOPN-----GRCALDGLGTTQLLPTAICSFGRGRINQKVSNGENHVMNM 303
Db 206 QSPCVSPKTTDPEEGFPRGLGACTLLGSPQHSPTSPRASVTEESWLGARSRPASPCNK 265
Qy 304 QVTNING--TPDTP-TGDVPAPLGTDFDS--GKLFGLVLSQRDHDNACRSHDAVATNSA 357
Db 266 RYKSLNGRQPPYPSPHSPPTSPHSGPRSVTDDSLGNTTQ-----YTSSAIVAAINA 318
Qy 358 KFTPKLGAIOIGTWEEDDVIH-----NQP-----TKFTPVCLFENEGFNQWTLNYSGA 406
Db 319 LTLD--SSLDLG---DGVFVKSRTTLEQPPSVALKVEPVG----- 354
Qy 407 LTLNMGGLAPPVAPTFPEGEQILFPRSHIPKGGVADPVIDCLLPQEWI OHLYESAPS--- 463
Db 355 --EDLGSPPPPADAFEDYSSP--QHI-RKGGFCQOYL--AVP-----QHPYQWAKPKPLS 403
Qy 464 -----QSDVALIRFTNPTGRVLFPEAKL-HRSGYIT-----VANTGSRPIVVP 505
Db 404 PTSYMSPTLPALDWQLPSSHGPGYELRIEVQPKSHHRAHYETEGSRGAVKASAGGHP1-VQ 462
Qy 506 ANGYFRDFTWVNOFYSLAPMGTGNGR 531
Db 463 LHGYLENEPLMLQLF-----IGTADDR 484

RESULT 12
US-10-116-275-182
; Sequence 182, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; TITLE OF INVENTION: Compositions targeting Peyer's Patches and M Cell Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116.275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 182
; LENGTH: 716
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-182

Query Match 3.9%; Score 110; DB 15; Length 716;
Best Local Similarity 21.5%; Pred. No. 0.48;
Matches 109; Conservative 62; Mismatches 201; Indels 134; Gaps 26;
Qy 111 GNAFTAGKLVAFAVPPH--FPLENISPGQITMPPHVI-----DVRTLEPVLPLPDVNR 163
Db 28 GPAPRAGGTMKSAEEHYGASSNVSPALPLPTAHSTLPAPCHNLQSTPGIIPADHPS 87
Qy 164 NF-----FHYNQ-----NEPRMLVAMLYTPLRSNGSGDDVFTVSCRVL 203
Db 88 GYGAALDGCAGYFLUSSGHTPDGAPALSPRIITSCGLYHNNNOFFHDEVEDVLPS 147
Qy 204 TRPSDFDNFVLPPTLESKTKPFTL-PILTIGELT-----NSRFPVPIDELYTSPN 254
Db 148 SKRSP--STATLSLPSLEAYRDPSCLSPASSLSRSCNSEASSYESNYSPYASPTQSPW 205

Wed Jun 2 09:13:33 2004

us-09-926-799-8.rapb

QY 246 IDELYTSPNESLVOPQNGRCALDGELOQTOLLPTAICSPRGRINOKV-----SG 296
Db 609 V-TVYASIR-----QASPIRLASVTALIESVNGKTVTLELLDNGAG 649
QY 297 ENHVNMQVNTINGTTPDPTGVPAPLGTDFSGKLF---GVLSQRDHDNACRSHDAVIA 353
Db 650 ADATKNDGVYSRFFTAFDANG-----RYSVKIWAALGGVTSRDROR-----688
QY 354 TNSAKFTPKLGAIQIGTWEED-DVHINOPTKFTPVGLFENEGFNOWTLPNYSGA--LTLN 410
Db 689 ----AAPPKNRMYIDGMIEDGEVRNMPRPETSVQDKQLCFSETS-----SGGSFVATN 740
QY 411 MGLAPVAPTFPGEQILFFRSHIPKGGVADPVIDCLLPQEWIOHLYOESAPSDVALI 470
Db 741 VPAAPIPDLFPCCQITDLKASIQQN-----LVNLTW-----TAPGD-----778
QY 471 RFTNPTGRV-----LFEAKLHRSYITVANTGSRPIVVPANGYERFDTWVNOFY 520
Db 779 ---DYDHGRASNYIIRMSTSIVDLRDHFNTSLQVNTTGLIPKEASSEIFEELGGNTF- 834
QY 521 SLAPMGTTGNG 530
Db 835 -----GNG 837

Search completed: June 1, 2004, 14:04:42
Job time : 36.2005 secs

Db 609 V-TVYASIR-----QASPIRLASVTALIESVNGKTVTLELLDNGAG 649
QY 297 ENHVNMQVNTINGTTPDPTGVPAPLGTDFSGKLF---GVLSQRDHDNACRSHDAVIA 353
Db 650 ADATKNDGVYSRFFTAFDANG-----RYSVKIWAALGGVTSRDROR-----688
QY 354 TNSAKFTPKLGAIQIGTWEED-DVHINOPTKFTPVGLFENEGFNOWTLPNYSGA--LTLN 410
Db 689 ----AAPPKNRMYIDGMIEDGEVRNMPRPETSVQDKQLCFSETS-----SGGSFVATN 740
QY 411 MGLAPVAPTFPGEQILFFRSHIPKGGVADPVIDCLLPQEWIOHLYOESAPSDVALI 470
Db 741 VPAAPIPDLFPCCQITDLKASIQQN-----LVNLTW-----TAPGD-----778
QY 471 RFTNPTGRV-----LFEAKLHRSYITVANTGSRPIVVPANGYERFDTWVNOFY 520
Db 779 ---DYDHGRASNYIIRMSTSIVDLRDHFNTSLQVNTTGLIPKEASSEIFEELGGNTF- 834
QY 521 SLAPMGTTGNG 530
Db 835 -----GNG 837

RESULT 15
US-10-369-214-132
; Sequence 132, Application US/10369214
; Publication No. US2003023037A1
; GENERAL INFORMATION:
; APPLICANT: Groot, Pieter C.
; APPLICANT: Berghenhegouwen van, Bram J.
; APPLICANT: Oosterhout van, Antoon J.M.
; TITLE OF INVENTION: Genes involved in immune related responses observed
; TITLE OF INVENTION: with asthma
; FILE REFERENCE: P53837US00
; CURRENT APPLICATION NUMBER: US/10/369,214
; CURRENT FILING DATE: 2003-02-15
; PRIOR APPLICATION NUMBER: EP 00202867.8
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: PCT/NL01/00610
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 132
; LENGTH: 913
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(913)
; OTHER INFORMATION: /note="Calcium-activated chloride channels Gob-5"
US-10-369-214-132

Query Match 3.8%; Score 109; DB 15; Length 913;
Best Local Similarity 21.5%; Pred. No. 0.87;
Matches 118; Conservative 60; Mismatches 208; Indels 164; Gaps 25;
QY 31 PVAGASIAAAPTGGQNNIIDPWIRLNFVQAPNGEFTVSPRSPGVELNLELGPENLYLA 90
Db 402 PTDGSEIVLLTDGEDNTISSCFDL-----VKQSGAIIHTVALGPAAKELE 447
QY 91 HLSRMVYAGGVFVQVLLAG--NAFTAGKLVFAAVPPH-FPLE---NISPGQITWFPH 143
Db 448 QLSKMTGGLQTVSSDOVQVQNGVLVDAPALSSGNAIAQHSIQLESRGVNLQNNQ-WWNGS 506
QY 144 VIID-----VRTLEPVLLPLPD---VRNNFFHYNOQNEPRMLVAMLYTPLRSN 189
Db 507 VIVDSSVGKDTLFLITWTHPTIFIDWPSGVEQNGFILDITTK-----VAYLQVPGTAK 561
QY 190 -GGDDVFTVSCRVLTRPSDFDFNLYVPPTLESKTKPTLPILTIGELTN---SRFPVP 245
Db 562 VGFWKYSIQASSQTTLT-----TVTSRAASATLPIIVTPVKNKNTCKFPSP 608

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:34:56 ; Search time 6.46309 Seconds
(without alignments)
4310.252 Million cell updates/sec

Title: US-09-926-799-8

Perfect score: 2843

Sequence: 1 MKMASNDAPSDGAGLVLP.....VNQFSLAPMGTGNGRRRVQ 535

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1925	67.7	539	1	COAT LORDV
2	1174	41.3	546	1	COAT SOUV3
3	325.5	11.4	2344	1	POLN_RHDV
4	289	10.2	702	1	COAT SMSV1
5	284	10.0	668	1	COAT FCVF4
6	277.5	9.8	671	1	COAT FCVF9
7	273.5	9.6	703	1	COAT SMSV4
8	272.5	9.6	668	1	COAT FCVC6
9	262	9.2	2208	1	POLN_MANCV
10	142	5.0	2193	1	POLG_CX16T
11	137	4.8	2193	1	POLG_CX16G
12	130	4.6	2206	1	POLG_POL32
13	128	4.5	2194	1	POLG_HE701
14	125.5	4.4	2333	1	POLG_FMDV1
15	125.5	4.4	4705	1	FAT2_DROME
16	122	4.3	980	1	CLV1_ARATH
17	121.5	4.3	2332	1	POLG_FMDV0
18	120.5	4.2	2205	1	POLG_POL2W
19	120	4.2	1011	1	POLG_FMDVT
20	119.5	4.2	2336	1	POLG_FMDVZ
21	118	4.2	2193	1	POLG_HE71M
22	115	4.0	2332	1	POLG_FMDVA
23	114	4.0	2207	1	POLG_POL2L
24	112	3.9	2206	1	POLG_POL3L
25	109	3.8	2193	1	POLG_EC09H
26	109	3.8	2201	1	POLG_CXA9
27	108.5	3.8	808	1	POLG_HP4V3
28	108	3.8	2193	1	POLG_HE71B
29	107	3.7	788	1	BCSB_XANAC
30	106.5	3.7	1739	1	DOTL_HUMAN
31	106	3.7	943	1	NFCL_HUMAN
32	104	3.7	2184	1	POLG_EC01F
33	102.5	3.6	1698	1	41_DROME

34 101.5 3.6 852 1 POLG_HP4VC
35 101.5 3.6 2185 1 POLG_SVDVU
36 101.5 3.6 2196 1 POLG_EC05N
37 101.5 3.6 2206 1 POLG_POL1M
38 101.5 3.6 2230 1 POLG_HP4VS
39 100.5 3.5 604 1 RCO1_NEUCR
40 100.5 3.5 955 1 AGAB_VIBS7
41 100.5 3.5 2208 1 POLH_POL1M
42 100.5 3.5 2226 1 POLG_HP4V2
43 100.5 3.5 2226 1 POLG_HP4V4
44 100.5 3.5 2227 1 POLG_HP4VH
45 100.5 3.5 2227 1 POLG_HP4VL

ALIGNMENTS

RESULT 1

ID COAT LORDV STANDARD; PRT; 539 AA.
AC P54635;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Coat protein (Capsid protein).
OS Lordsdale virus (Human enteric calicivirus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=82658;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96005060; PubMed=7561776;
RA Dingle K.E., Lambden P.R., Caul E.O., Clarke I.N.;
RT "Human enteric Caliciviridae: the complete genome sequence and expression of virus-like particles from a genetic group II small round structured virus";
RT J. Gen. Virol. 76:2349-2355(1995).
CC - SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

CC EMBL; X86557; CAA60255.1; --
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 531 531 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 539 AA; 58775 MW; 585C63E7F2C5FD21 CRC64;

Query Match 67.7%; Score 1925; DB 1; Length 539;

Best Local Similarity 65.7%; Pred. No. 6.3e-137;

Matches 355; Conservative 79; Mismatches 96; Indels 10; Gaps 5;

QY 1 MKMASNDAPSDGAGLVLPPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60

Db 1 MKMASNDANPSDGSANLVPEVNEVMALEPVGNAIAAPVAGQNVIDPWIRNFFVQAP 60

QY 61 NGEFTVSPRNPGVEVLNLELGPBLNPLYAHLSRMYNGYAGVEVQVLLAGNAFTAGKLV 120

Db 61 GGEFTVSPRNAPGAILWSAPLGPLNPLYSLSRMYNGYAGGFEVQVLLAGNAFTAGKVI 120

QY 121 FAAPPPHPEPLNISPGQITMEPHVITDVLTLEPVLPLPDVNRNPFHYNQNEPRMRLVA 180

Db 121 FAAPPPNPFTEGLSPSQVTMPFHIIVRQLEPVLPLPDVNRNPFHYNQNDSTLKLIA 180

1	MMASKADPQSDAGSAGQVLPVNTADPLPMPVAGPTTAVATAGQVNMIDPVIWNVF	60
57	VOAPNGEFTVSPRNSPGEVLNLELGPALNPLAHLRSMYNGVAGGVEVOVLGNAFTA	116
61	VQSPQCEFTISPNTPGDIPLDQLGPHLNPFLSHLSQMTNGWGNMRVILLAGNAFTA	120
117	GKLVFAAAPPHPFPLENISPGQITWFFPHVIDVRLTLEPVLPLPDRNNFPHYNOQNEPRM	176
121	GKIIIVCCVPPGFTSSSLIAQATLFPHVIAVRLTLEPVLPLPDRNNFPHYNOQNEPRM	179
177	RLVAMLYTLPLRS-NGSGD-DVFTVSCRVLTRPSPDFDFNYLVLPPTLESKTKPFTLPLITI	234
180	RLVCMLYTLPLRTGGSGNSDSFWAGRVLTAPSSDFSFLVLPPTIEQKTRAFVFNPL	239
235	GELTNSRFPVPTDELTYSPNESLVVQPNQRCALDGELOQTQLLPTAICSFGRINQKV	294
240	QTLNSRFPVPSLQGMILSPDASQVQFQNGRCLIDGQLLTPATSGQLFRVRGKINQGA	299
295	SGENHVMNQVNTINGTPDTPGVPAPLGTDPDPFGKLVGVLSDRDHONACRSHDAV---	351
300	R-----TLNLTEVDGKPF-MAFDSPAPVGFDPF-GKCDWHMRISKTNNNTGSGDPMRSV	351
352	-IATNKAFTPKLGAIQ-----IGTWEDDDVHINOPTKFTPTVGLFENEFGNQW	398
352	SVQTNVQGFVPHUGSIQFDEVFNHPTGDIYIGIE-----WISQPS--TPG---TDINLW	401
399	TLPNYSGLTNNMGLAPPVAPTPPGSQILFFRSHIPLKGGVADP-VIDCLLPQEWIOLHY	457
402	EIPDYSSLSQANLAPPVFPFGFGEALVYVSAFEGPNNRSPNDVPCLLPQEIYTHFV	461
458	QESAPSDVALTRFTNPDTPGRVLFPAKLHRSYITVA--NTGSRPIVVPVANGYFRFTW	515
462	SEQAPTGMGAALLHYVDPDTRNLGFBKLYPGGYLTCVPNGVAGPQQLPLNGVFLVSW	521
516	VNOFYSLAPMGTGN-----GRRRV	534
522	VSRFYQLKXPGVGTASTARGRLGVRI	546
RESULT 3		
1	POLN RHDV STANDARD; PRT; 2344 AA.	
AC	P27410; (Rel. 23, Created)	
DT	01-AUG-1992 (Rel. 23, Last sequence update)	
DT	10-OCT-2003 (Rel. 42, Last annotation update)	
DE	Non-structural polyprotein [Contains: RNA-directed RNA polymerase (EC 2.7.7.48); Thiol protease P3C (EC 3.4.22.-); Helicase (2C like protein); Coat protein].	
OS	Rabbit hemorrhagic disease virus (RHDV).	
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae; Lagovirus.	
OX	NCBI_TaxID=11976;	
RN	SEQUENCE FROM N.A.	
RP	MEDLINE=91361557; PubMed=1840711;	
RA	Meyers G, Wirblich C, Thiel H.-J.;	
RT	"Rabbit hemorrhagic disease virus -- molecular cloning and nucleotide sequencing of a calicivirus genome.";	
RL	Virology 184:664-676(1991).	
CC	-I- FUNCTION: THE P3C CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR CLEAVAGE AT CERTAIN Q/G SITES IN THE POLYPROTEIN.	
CC	-I- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA) (N).	
CC	-I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
181	MLYTLRNGSGDDVFTVSCRVLTTRPSPDFDFNYLVLPPTLESKTKPFTLPLITIGELTNS	240
181	MLYTLRANNAGDDVFTVSCRVLTTRPSPDFDFNYLVLPPTLESKTKPFTVPLVTEEMNS	240
241	RFPVIDELYTSPNESLVVQPNQRCALDGELOQTQLLPTAICSFGRINQKVSNGNHV	300
241	RFPVPLEKLYTGFSSAFVQPNQRCALDGELOQTQLLPTAICSFGRINQKVSNGNHV	298
301	WNNQVNTINGTPDTPGVPAPLGTDPDPFGKLVGVLSDRDHONACRSHDAVIATNSAKF	359
299	YTMNLASQNSWYDPTPEIPAPLGTDPDPFGKLVGVLSDRDHONACRSHDAVIATNSAKF	358
360	TPKLGAIQIGTWEDDHHNOPTKFTPKVGLF-----ENEGSNOWTLNYSGLTNNMGL	413
359	TPKLGSVQFTTNDNFQAGQNTKFTPVGVIOGDHQNKP-QQWLSLNYSGRTGHNHVL	417
414	APVAPTFPGEQILFFRSHIPLKGGVADPVIDCLLPQEWIOLHYQESAPSDVALTRFT	473
418	APVAPTFPGEQILFFRSTMPGCGYPNMNLDCLLPQEWIOLHYQESAPSDVALTRFT	477
474	NPDTGRVLFPAKLHRSYITVAQTSRPIVVPVANGYFRFTWNOFYSLAPMGTGNRRR	533
478	NPDTGRVLFPAKLHRSYITVAQTSRPIVVPVANGYFRFTWNOFYSLAPMGTGNRRR	537
RESULT 2		
1	COAT SOUV3 STANDARD; PRT; 546 AA.	
AC	Q04542; (Rel. 30, Created)	
DT	01-OCT-1994 (Rel. 30, Last sequence update)	
DT	01-OCT-1994 (Rel. 30, Last annotation update)	
DE	Coat protein (Capsid protein).	
OS	Southern virus (serotype 3).	
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae; Norovirus.	
OX	NCBI_TaxID=37129;	
RN	SEQUENCE FROM N.A.	
RP	MEDLINE=93142023; PubMed=8380940;	
RA	Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;	
RT	"Sequence and genome organization of a human small round-structured (Norwalk-like) virus.";	
RL	Science 259:516-519(1993).	
CC	-I- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; L07418; AAA92984.1; --	
DR	PIR; B37491; B37491.	
DR	InterPro; IPR004005; Calici_coat.	
DR	InterPro; IPR008975; Viral_cap_coat.	
DR	Pfam; PF00915; Calici_coat; 1.	
KW	Coat protein; Glycoprotein.	
FT	CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).	
SQ	SEQUENCE 546 AA; 58774 MW; C82B2A85AD4B05EA CRC64;	
Query Match		
Best Local Similarity 46.2%; Pred. No. 1.5e-80;		
Matches 263; Conservative 82; Mismatches 172; Indels 50; Gaps 16;		
1	MMASNDAAASNDGAAG--LVPEANN-ETMALEPVAGASIAAPLTGQNNIDPWRILNF	56

CC EMBL; M67473; AAA47285.1; --
 DR F1R; A41039; RRRWRH.
 DR PDB; 1KHV; 16-JAN-02.
 DR PDB; 1KHV; 16-JAN-02.
 DR MEROPS; C24.001; --
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR000317; Peptidase_C24.
 DR InterPro; IPR000605; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF03510; Peptidase_C24; 1.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00916; 2CENDOPTASE.
 DR PRINTS; PR00918; CALICIVIRUSNS.
 DR SMART; SM00382; AAA; 1.
 KW Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;
 KW Polyprotein; Helicase; ATP-binding; Coat protein; 3D-structure.
 FT CHAIN ? 7508 7707 HELICASE (P2C-LIKE).
 FT CHAIN 1109 1251 PROTEASE P3C.
 FT CHAIN 71497 71625 RNA-DIRECTED RNA POLYMERASE.
 FT CHAIN ? 2344 COAT PROTEIN.
 FT NP BIND 522 529 ATP (POTENTIAL).
 FT ACT_SITE 1212 1212 PROTEASE (POTENTIAL).
 FT ACT_SITE 1227 1227 PROTEASE (POTENTIAL).
 SQ SEQUENCE 2344 AA; 257066 MW; 1454C24F81E9212 CRC64;
 Query Match 11.4%; Score 325.5; DB 1; Length 2344;
 Best Local Similarity 26.3%; Pred. No. 5.2e-16;
 Matches 152; Conservative 73; Mismatches 227; Indels 125; Gaps 30;
 QY 8 AAPNDGAAGLVPEAN---NETMALEP-----VAGASIAAPLTGQNNI 47
 DB 1772 AAPQGE-AAGTATTASVPGTTGMDPGVAVTTSVITAENSASITAGIGGP-PQOVQ 1829
 QY 48 IDPWIRLNFVQAPNGEFTVSPNSPCEVLNLELGPENLPYLHLRMYNGVAGGVEQV 107
 DB 1830 QETM-RTNFYY---NDVFTWSVADAPGSILYTVQSPNNPFTAVLSQMYAGWAGGQFRF 1886
 QY 108 LLAGNAFTAGKLVFAAVPPHPLENISPG-QITMPHPHVIDVRLPVLPLPDVNRNFF 166
 DB 1887 IVAGSGVFGRLVRAVIPPGL---EIGPGLVRFQPHVVIDARSLEPVITTPDLRPNNY 1943
 QY 167 HYNQNEPRM---RLVAMLYTPILRSNGSGDDVFTVSCRVLTRPSPDPDFNLYVPPTLESKT 224
 DB 1944 H--PTGDPGLVPTLVLSVYNNL--INPFGGSTSAIQVTVEPSPDEFEFVIRAPS--SKT 1998
 QY 225 ----KP---FTPLTIGLTNSRFPVPIDELTSNLSV-VQPNQ---RCALDGELO 273
 DB 1999 VDSISPAGLLTTPVLT-VGVNDNRW-----NGQIVGLQVPYGGFSTCNHRWNLN 2046
 QY 274 GTTQLLPTA----ICSFGRINQKVSNGENHVMNQVNTINGTPFDTPGVAPLGTPDFS 329
 DB 2047 GSTYGHSSPRFGDIDHRRGSASTSGSNATNVLQFTYANAGSAIDNPISQV-APDGPDM 2105
 QY 330 GKLFVGLSQRDHDNACRSHDAVIATNSAKFTPKLGAIGT-WEEDDVHINOPT-KFTPV 387
 DB 2106 FVPF-----NGPGIPAGWVGFGAIWNSGAPNVTVOAYEL 2143
 QY 388 GLFENEGFNQWTLPNVSGALTNMGLAPVAPFTFGQILPFRSHIPLKGGVADVIDCL 447
 DB 2144 GEATGAPGNLQPTNTTSGATVAKSYAVVTGTAQNPAGLF-----VMASGIISTPNASAI 2199
 QY 448 L----PQEWIOHLYQSSAQSDVALIRFTNPDTRGVLPFAKLHRSYI-----TV 494
 DB 2200 TYTPQPDRIV-----TTPGTAAAPVGNKTP-----IMPVSVRRTGDNVATAGSANGTQ 2249

QY 495 ANTGRPIVW-----PANGYF-----RDTWVNOFYS 521
 DB 2250 YGTGSQPLPFTVITGLSNLYSALMPEGQFFVWQLTFFAS 2286
 RESULT 4
 COAT_SMSV1 STANDARD; PRT; 702 AA.
 ID AC P36284;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Coat protein (Capsid protein).
 OS San Miguel sea lion virus (serotype 1) (SMSV 1).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OC NCBI_TaxID=36406;
 RN [1]_TaxID=36406;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92410750; PubMed=1529644;
 RA Neill J.D.;
 RT "Nucleotide sequence of the capsid protein gene of two serotypes of
 RT San Miguel sea lion virus: identification of conserved and non-
 RT conserved amino acid sequences among calicivirus capsid proteins.";
 RL Virus Res. 24:211-222(1992).
 CC -|- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC EMBL; M87481; AAA16217.1; --
 DR F1R; A48562; A48562.
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_Cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 481 481 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 493 493 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 545 545 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 702 AA; 77850 MW; E6E5A58523DEB3D7 CRC64;
 Query Match 10.2%; Score 289; DB 1; Length 702;
 Best Local Similarity 25.4%; Pred. No. 5.5e-14;
 Matches 128; Conservative 62; Mismatches 190; Indels 124; Gaps 23;
 QY 8 AAPNDGAAGLVPEANNTMALEPVAGAS-----IAAPLTGQNNIIDPWLRLNFVQAPNGE 63
 DB 151 AESDGFAGADIVTEQGTVVOQOPVPAQSAALTLLAASGTKVDCEWTTFFSYHTAVNWS 210
 QY 64 FTVSPRNPGEVLLNLELGPENLPYLHLRMYNGVAGGVEVQVLLAGNAFTAGKLVFAA 123
 DB 211 TT-----EAQKILFGRALSPENLPYLHLRMYSSYSTWSGIDVRFTVSGSGVFGKLAALI 266
 QY 124 VPPHPLENISPGQITMPHPHVIDVRLPVLPLPDVNRNFFHYNQNEPRMLVAMLY 183
 DB 267 VPP--GIEPVESPTMLQYPHVFLDARQTEPVITFDIKRTLYH-SMDDTDTTLLVIMY 323
 QY 184 TPL-RSNGSGDDVFTVSCRVLTRPSPDPDFNLYVPPTLESKTKPTLPLTIGELTNSRF 242
 DB 324 NELINPYQSEPKSCSCTIVETRPSPDFTSLKPP--GSLKKGSIPLDLP--NSRH 379
 QY 243 PVPIDELTSPNESLWQPO-----NGRCALDGELOG----- 274
 DB 380 --WMGNRWSTIDGVFVQPRVFNQSRHDFDFTTTTGWSTPPYVPIEVLTLEKLDRGQYFK 437
 QY 275 -----TTQLLPTAICSFRGRINQKVSNGENHVMNQVNTINGTPF--- 313

438	VTDTEKSLVPLPGDWPTT--IPTAMTASNGNYDYTV-----BYRITN-NGTHFKGF	488
	Db	
314	-----DPTGDVPAPLCTPDFSKLFGVLRSQRDHDNACR-----SHDAVI--	352
	Qy	
489	YIMGNLTTKVKGSNLGE-----TQTSRTLFASVGNYKDQNTINPTHKITSNSLVVYD	542
	Db	
353	ATNSAKFTPKLGAIQIGTWEEDDVHIINQP-TKFTPVGLFENEGFNQWTLPNYSGALTLNM	411
	Qy	
543	ANNVSAATAK-----TTTWHSTMSHLGVLVDESEVGSdstKVVRIATLPE--AFT-NG	593
	Db	
412	GLAPPVAPTFPGEQILPFRSHIPL	435
	Qy	
594	G-----NFP-----VFFTNKIQI	606
	Db	

```

RESULT 5
COAT_FCVF4
ID COAT_FCVF4 STANDARD; PRT; 668 AA.
AC F27405;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Coat protein (Capsid protein).
OS Feline calicivirus (strain Japanese F4) (FCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesiviruses.
OC NCBI_TaxID=11980;
CX [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=91306470; PubMed=18535578;
RA TONYA Y., TANIGUCHI Y., TAKAHASHI E., UTAGAWA E., TAKEDA N.,
RA MIYAMURA K., YAMAZAKI S., MIKAMI T.;
RT "Sequence analysis of the 3'-end of feline calicivirus genome.";
RL Virology 183:810-814(1991).
CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

QY		185	PL-----RNSGGDDVFTVSCRVLTRPSDPDFENYLVPPTLESKTKPFTLPILTIGELT	238
Dd		294	LINPYANDSNSSCIV-----TVETKPGDFKHLLKPPG-----SVLTHGSIP	338
QY		239	NSRRFP-----VPIDELYTPSNESLVQP-----QNGRCALDGELOG--TTQLLPATICSFRG	288
Dd		339	SDLIPKSSSLMTGNRYWTDIIDFVIRPEVFQANRHFDNQETAGWSTPRRPITI-----	393
QY		289	RINQKVSGENVHWNMQVTNINGTP-----FDPCTGDVPAPLG-----	324
Dd		394	TISEXNKGSKIGVGATDYIIIFGIQDGWPDTIADKLIPAGDYSITTEGNDIKTAQAYDT	453
QY		325	-----TFDFSGLFGVLSQRDHONACRSHDAVIAT-----NSAKFTPCLKAIQIGTWEE	373
Dd		454	AAVVKNNTINFRGMWYCISLRAGWDKKISHTAFITTAIRDGNEIKPSNTIDMTKLAVIQ-	512
QY		374	DVHNINOPTK-----TFPVCLFENECSFNQTWLTPNYSGALTLMNLGLAPPVPAPTFPG	423
Dd		513	-DTHVEQEVOVSDDTLALLGYTGIG-EBAIGNRDVRVIS-----VLPAGARGG	561
QY		424	EQILPFERSHIP L 435	
Dd		562	NHEIFYNKS I K L 573	
 RESULT 6 COAT_FCVF9 STANDARD; PRT; 671 AA.				
ID	COAT_FCVF9	PRT;	671 AA.	
AC	P27406;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	01-JUN-1994 (Rel. 29, Last annotation update)			
DE	Coat protein (Capsid protein).			
GN	CpB76			
OS	Feline calicivirus (strain F9) (FCV).			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;			
OC	Vesivirus.			
OX	NCBI_Taxid=11981;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92410623; PubMed=1529544;			
RA	Carter M.J., Milton I.D., Meanger J., Bennett M., Gaskell R.M.,			
RA	Turner P.C.;			
RT	"The complete nucleotide sequence of a feline calicivirus.";			
RL	Virology 190:443-448(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=923339673; PubMed=1633955;			
RA	Meanger J., Carter M.J., Gaskell R.M., Turner P.C.;			
RT	"Cloning and sequence determination of the feline calicivirus strain			
RT	F9.";			
RL	Biochem. Soc. Trans. 20:26S-26S(1992).			
CC	-1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	----			
DR	EMBL; M86379; AAA79327.1; -			
DR	EMBL; Z11536; CAA77636.1; -			
DR	PIR; B43382; VCNWF9.			
DR	InterPro; IPRO04005; Calici_coat.			
DR	InterPro; IPRO08975; Viral_Cap_coat.			
DR	Pfam; PF00915; Calici_coat; 1.			
DR	Coat protein; Glycoprotein.			
KW	CARBOHYD 177 177			
FT	CARBOHYD 304 304			
FT	CARBOHYD 439 439			
FT	(POTENTIAL). . . (POTENTIAL). . .			
FT	(POTENTIAL). . . (POTENTIAL). . .			

FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 671 AA; 73441 MW; 33BEE86D8370D5B5 CRC64;

Query Match
 Best Local Similarity 9.8%; Score 277.5; DB 1; Length 671;
 Matches 76; Conservative 38; Mismatches 97; Indels 31; Gaps 9;

QY 30 EPVAGASIAAPITGQNNIIDPW-IRLNFVQAPNGEFTVSPRNSPGEVLLNLGLPGLNPPY 88
 DB 145 EPSAQWSTAADWATGKSDVSEWAEFPSPHTSVNW-----STSETQKILFKQSLGILLNPPY 200
 QY 89 LAHLRSMYNGYAGGVEVQVLLAGNAFTAGKLVFAAPVPPHPIENISPGQIT---MPPHVI 145
 DB 201 LEHLAKLYVWAGSIEVRSISGSGVFGKLAIVVPP-----GVDPPQSTSMQLQPHVL 255
 QY 146 IDVRLTEPVLPLDVRNNFFHYNQNEPRMELVAMLYTPLRSNGSGDDVFTVSC--RVL 203
 DB 256 FPARQVEPVIFCLPDLRSTLYHLMSTD-TTSLVIMVYNDL-INPYANDANGSGGIIVYE 313
 QY 204 TRPSPDFDNVLPPTLESKTPFTLPILTIGELTNSRP-----VPIDELTYSPNESLVV 259
 DB 314 TRPGPDFKFLKPPG-----SMUTHGSIPSDLIPKTSLSLWIGNRYWSDITDFVI 363
 QY 260 QP 261
 DB 364 RP 365

RESULT 7
 COAT_SMSV4
 ID COAT SMSV4 STANDARD; PRT; 703 AA.
 AC P36285;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Coat protein (Capsid protein).
 OS San Miguel sea lion virus (serotype 4) (SMSV 4).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OX NCBI_TaxID=36407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92410750; PubMed=1529644;
 RA Neill J.D.;
 RT "Nucleotide sequence of the capsid protein gene of two serotypes of
 RT San Miguel sea lion virus: identification of conserved and non-
 RT conserved amino acid sequences among calicivirus capsid proteins.";
 RL Virus Res. 24:211-222(1992).
 CC -|- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M87482; AAA16220.1; -.
 DR PIR; C48562; C48562.
 DR InterPro; IPR004005; Calici coat.
 DR Pfam; PF00915; Calici coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 703 AA; 77721 MW; C5DAD8223B261073 CRC64;

Query Match
 9.6%; Score 273.5; DB 1; Length 703;

Best Local Similarity 28.8%; Pred. No. 8e-13;
 Matches 97; Conservative 51; Mismatches 150; Indels 39; Gaps 15;

QY 8 AAPSDNGAAGLVPEANNETM-----ALEPVAGASIAAPITGQNNIIDPWIRLNFVQAPNGE 63
 DB 151 AESDQFGSAEIVTEGQTVVQOOPAPAPATLATATATGK-SVEQEMWTFPSYHTSINW 209
 QY 64 FTVSPRNSPGEVLLNLGLPGLNPPYLAHLRSMYNGYAGGVEVQVLLAGNAFTAGKLVFAA 123
 DB 210 STV---ESQKILYQALNPSINPYLDHIAKLYSTWSGIDVRFTVSGSGVFGKLAALL 266
 QY 124 VPPHF-PLENISPGQITMPPHVIIDVRLTEPVLPLDVRNNFFHYNQNEPRMELVAML 182
 DB 267 VPPGVEPIESVS---MLOYPHVLFDPARQTEPVIFTPDIRKTLFHSMDTDT-TTKLVINP 322
 QY 183 YPLRSNGSGDDVFTVSCRVLTRPSPDFDNVLPPTLESKTPFTLPILTIGELTNSRF 242
 DB 323 Y-----ENGV-ENKTKTSITVETPSADTFALLKPP--GSLIKHSIP-----SDLIPRS 371
 QY 243 VPVIDELTYSPNESLVVQPO-----NGRCALDGELOG--TTQLLPTAICSFGRINQKVG 296
 DB 372 AHMGNRMWSTISGFSQPRVFQSNRHFDFDSTTTGWSPTYVYPIEI-----KIQGKVG 426
 QY 297 EN---HVNMQVTNINGTTPDPDGDVPAPLCTPDPFSG 330
 DB 427 NKMFWHVIDTKALVPGIP-DGWPDTTIPDETATNG 462

RESULT 8
 COAT_FVCV6
 ID COAT FVCV6 STANDARD; PRT; 668 AA.
 AC P27404;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Coat protein (Capsid protein).
 OS Feline calicivirus (strain CFI/68 FIV) (FCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OX NCBI_TaxID=11979;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=91374597; PubMed=1716692;
 RA Neill J.D.; Reardon I.M.; Heinrichson R.L.;
 RT "Nucleotide sequence and expression of the capsid protein gene of
 RT feline calicivirus".
 RL J. Virol. 65:5440-5447(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Neill J.D.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 CC -|- PTM: The N-terminus is blocked.
 CC -|- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M32819; AAA42925.1; -.
 DR EMBL; U13992; AAC13993.1; -.
 DR PIR; A40507; VCMWFF.
 DR InterPro; IPR004005; Calici coat.
 DR InterPro; IPR008975; Viral cap coat.
 DR Pfam; PF00915; Calici coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 459 459 N-LINKED (GLNAC. . .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLNAC. . .) (POTENTIAL).
SQ SEQUENCE 668 AA; 73550 MW; 9E52312108D441 CRC64;

Query Match 9.6%; Score 272.5; DB 1; Length 668;
Best Local Similarity 31.3%; Pred. No. 8.9e-13;
Matches 77; Conservative 37; Mismatches 93; Indels 39; Gaps 9;

QY 30 EPVAGASIAAPLTGQNNIIDPW-IRLNFVQAPNGEFTVSPRNSPGEVLLNLELGPENPY 88
DB 145 EFNAQWSTADMATGKSDVSEWAFSEHTSNV-----STSETQKILFKOSLGLPLNLY 200
QY 89 LAHLSRMYNGYAGGVEQVLLAGNAFTAGKLVFAAVPHFPLENISPQGIT---MPPHVI 145
DB 201 LTHLAKLYVANGSDVDRFSISGSGVFGKLAIVVPP-----GIDPVQSTMLQYPHVL 255
QY 146 IDVRTLEPVLPLPDPVRNNFFHYNQNEPRMLVAMLYTPL-----RSNSGDDVFTVS 199
DB 256 FDAQVEPVIFSIPDLRSTLYHLSMDT-TTSLIVMVYNDLINPYANDNSGCVI----- 310
QY 200 CRVLRPSDPDFNLYVPPTLESKTPFTLPILTIGELTNSRPP---VPIDELYTSPNE 255
DB 311 -TVETKPGDPFKFHLKPPG-----SMLTHGSPDSLIPKSSSLWIGNRFSWSDIT 359
QY 256 SLVVQP 261
DB 360 DFVIRP 365

RESULT 9
POLN MANCV STANDARD; PRT; 2208 AA.
ID -POLN MANCV
AC Q63014;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein (Contains: RNA-dinected RNA polymerase
DE (EC 2.7.7.48); Thiol protease 3C (EC 3.4.22.-); Helicase (2C like
DE protein); Coat protein).
OS Manchester virus (Human enteric calicivirus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Sapovirus.
OX NCBI_TaxID=82659;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95390791; PubMed=7661689;
RA Liu B.L., Clarke I.N., Caul E.O., Lambden P.R.;
RT "Human enteric caliciviruses have a unique genome structure and are
RT distinct from the Norwalk-like viruses."
RL Arch. Virol. 140:1345-1356(1995).
CC -!- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).
CC -!- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR
CC THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X86560; CAA60262.1; --
CC MEROPS; C24.0PW; --
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR000317; peptidase_C24.

DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_F3D.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR InterPro; IPR008975; Viral_cap_coat.
DR pfam; PF00915; Calici_coat; 1.
DR pfam; PF03510; Peptidase_C24; 1.
DR pfam; PF00680; RNA_dep_RNA_pol; 1.
DR pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; 2CENDOPTASE.
DR PRINTS; PR00918; CALICIVIRUSNS.
DR Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;
KW Thiol protease; Helicase; ATP-binding; Coat protein.
FT CHAIN ? ? HELICASE (P2C-LIKE).
FT CHAIN ? ? PROTEASE (P3C).
FT CHAIN ? ? RNA-DIRECTED RNA POLYMERASE.
FT NP_BIND 408 415 ATP (POTENTIAL).
FT ACT_SITE 1097 1097 PROTEASE (BY SIMILARITY).
FT ACT_SITE 1112 1112 PROTEASE (BY SIMILARITY).
SQ SEQUENCE 2208 AA; 242736 MW; 3E299D5BA20E45DC CRC64;

Query Match 9.2%; Score 262; DB 1; Length 2208;
Best Local Similarity 21.0%; Pred. No. 2.8e-11;
Matches 133; Conservative 73; Mismatches 251; Indels 176; Gaps 23;

QY 2 KMASNDAAAPSDGAGLVPEANNETWALEPFAVAGASIAAPLTGQNNIIDPWIRLNFVQAPN 61
DB 1647 EMEGNSNP-----PKQSNPMVVDP---PGTTGTTSHVVVANP-----EQPN 1688
QY 62 G-----EFTVSPRNSPGEVLLNLELGPENLYLAH 91
DB 1689 GAAQRLLEAVATGAIQSNVPEAIRNCFAVETFAWDRMPTGTLGSLSHFNINPTSH 1748
QY 92 LSRMYNGYAGGVEQVLLAGNAFTAGKLVFAAVPHFPLENISPQGIT---MPPHVIDV 148
DB 1749 LSGWAGGSGFEVRLSISGSGVFAAGLIAVIPP---GVDPSSIRDPGVLPFAFVDA 1803
QY 149 RTLEPVLPLPDPVRNNFFHYNQNEPRMLVAMLYTPLRNSGDDVFTVSCRVLRPS 208
DB 1804 RITEFVSPMIDPRAVDYHRMDGAEPTCSLGFVYQPLLNPFFSTAVSTCWVSEYKPG 1863
QY 209 DFDFNYLVP--TLESKTKPFTLPILTIGELTNSRPPVIDELYTSPNESLVQPOQNGRC 266
DB 1864 DFDFCLLRPPGQOMENGVSPEGILLPRGLYSRGNRVGLVGMILVAEHKQVNRHFNNS 1923
QY 267 ALDGELOQTTLPTAICSFGRINQKVSNGENHVNQVNNINGTTPDPTGDVPAPLGP 326
DB 1924 VTFGHSAPVNPMAAEIVT---NQAHSSTSRHAW-LSIGAQNKGPLFP---GIP 1969
QY 327 DFSGKGLFVLSQRDHDAACRSHDAVIATNSAKFTPKLCAI--QIGTWEEDDVHINQP--- 381
DB 1970 N-----HPFDSCASTVVGAMDTSLGRRPSTGCGPAISFQNGNGDVYENDTPSV 2017
QY 382 -----TKFTPVGL-----FENEGFN-----QWTLPNYSGLAT 408
DB 2018 MFATYDPLTSGTGVALTNSINPASLALVRISNNDFTSGFANDKNVVVQMSWMTYGNQ 2077
QY 409 LNMGLAPPVAPT-----FPGEQLPFRSHIPLKGGVADPVIDCL 447
DB 2078 IR-GQVTPMSGTYTFTTGANTLVLMOERMUSYDGHQAILYSSQI----- 2122
QY 448 LPOEWIOLHYQESAPQSQSDVALIRFNTPDTRVLFPEAKLHRSGYITVANTGSRPIVSPAN 507
DB 2123 ---ERTAEYFQNDIVNIPENSMVAVF-NVETNSASFOIGIRPDGVMYTG--GSIGNVPLE 2176
QY 508 GYFRFDTWNQFYSLAPMG-----TGN-GRRR 533
DB 2177 PETRF-----QYVGILPLSAALSGPSGNMGRAR 2204

RESULT 10
POLG_CX16T

ID AC 09QF31; POLG CX16T STANDARD; PRT; 2193 AA.

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VPG (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].

DE Cocksackievirus A16 (strain Tainan/5079/98).

OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.

OC NCBI_TaxID=231417;

OX [1]

RN SEQUENCE FROM N.A.

RX MEDLINE=21427116; PubMed=11536241;

RA Yan J.-J., Su I.-J., Chen P.-F., Liu C.-C., Yu C.-K., Wang J.-R.;

RT "Complete genome analysis of enterovirus 71 isolated from an outbreak in Taiwan and rapid identification of enterovirus 71 and coxsackievirus A16 by RT-PCR.";

RL J. Med. Virol. 65:331-339 (2001).

CC -!- FUNCTION: It is thought that the P2C protein attaches to vesicular membranes and is associated with viral RNA synthesis.

CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.

CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-[Gly bond in the poliovirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.

CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA) (N).

CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.

CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.

CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AF177911; AAD55085.1; ..

DR HSP; P03300; IPOV.

DR InterPro; IPR003593; AAA_Atpase.

DR InterPro; IPR004004; Calici_pol_hel.

DR InterPro; IPR009003; Cys_Ser_trypsin.

DR InterPro; IPR000199; Pept_3C_picorn.

DR InterPro; IPR000081; Peptidase_C3.

DR InterPro; IPR003138; Pico_P1A.

DR InterPro; IPR002527; Pico_P2B.

DR InterPro; IPR001676; Rhv.

DR InterPro; IPR006005; RNA_helicase.

DR InterPro; IPR007095; RNA_pol_DS_PS.

DR InterPro; IPR001205; RNA_pol_P3D.

DR InterPro; IPR007094; RNA_pol_PSVir.

DR InterPro; IPR008975; Viral_cap_coat.

DR Pfam; PF00548; Cys-protease-3C; 1.

DR Pfam; PF02226; Pico_P1A; 1.

DR Pfam; PF00947; Pico_P2A; 1.

DR Pfam; PF01552; Pico_P2B; 1.

DR Pfam; PF00073; rhv_3.

DR Pfam; PF00680; RNA_dep_RNA_pol; 1.

DR Pfam; PF00910; RNA_helicase; 1.

DR PRINTS; PR00918; CALICIVIRUS.

DR ProDom; PD001125; Cys_protease_3C; 1.

DR ProDom; PD001306; Pico_P2A; 1.

DR ProDom; PD001274; Pico_P2B; 1.

DR SMART; SM00382; AAA; 1.

DR Polyprotein; Coat protein; Core protein; Transferase;

KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;

KW Lipoprotein.

FT CHAIN 2 69 COAT PROTEIN VP4.

FT CHAIN 70 323 COAT PROTEIN VP2.

FT CHAIN 324 565 COAT PROTEIN VP3.

FT CHAIN 566 862 COAT PROTEIN VP1.

FT CHAIN 863 1012 CORE PROTEIN P2A.

FT CHAIN 1013 1111 CORE PROTEIN P2B.

FT CHAIN 1112 1440 CORE PROTEIN P2C.

FT CHAIN 1441 1526 CORE PROTEIN P3A.

FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VPG.

FT CHAIN 1549 1731 PICORNAIN 3C.

FT CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE.

FT LIPID 2 N-myristoyl glycine (in host) (By similarity).

FT ACT_SITE 1695 PROTEASE (POTENTIAL).

FT ACT_SITE 1709 PROTEASE (POTENTIAL).

SQ SEQUENCE 2193 AA; 243182 MM; 927839DB58F61E7F CRC64;

Query Match 5.0%; Score 142; DB 1; Length 2193;

Best Local Similarity 22.7%; Pred. No. 0.028;

Matches 127; Conservative 60; Mismatches 223; Indels 150; Gaps 31;

QY 1 MKMASNDRAAPSNDG-AAGLVP-----EANNETMALEPVAGASIAAPLTGQNNI 47

DB 329 LKPGTNQFLTTDDGVSAFILPGFHPTPIHIPGEVHN---LLEICRVETILEVNNLKTN 385

QY 48 IDPMRLANF---VOAPNGEFTVSPRNSPGEVLLNLELGPENLPYLAHSLRMVNGVAGVE 104

DB 386 TTFMQRLCFPVSVQSKTGEELCAAFADPFORD-----GPMQSTILGQLCRYTQWSSGLE 439

QY 105 VOVLAGNAFTAGKLVFAAVPPHPPLENISQGIT--MPHHVIDVRLTEPVLPLPDVR 162

DB 440 VTFEAGSFMATGKMLIAYTPPG---GNVPADRIATMLGTHVWDGLQSSVTLVVPMTS 496

QY 163 NNFHYNQNEPRMPLVAMLYTPLASNGSGDDVFTVSCRVLTRPSDFDPFNVLPPTLES 222

DB 497 NT--HY-----RAHAR-AGYFDY-----511

QY 223 KTKPPTLPILTIGELTNSRFPVPIDELYTSPNESLVQPN---GRCALDGLQGTQJLL 279

DB 512 ----YTTGIIIIWYQTN--YVVPICAPTAYIVALAAQDNFTMKLCKDTEIDQTNITQ 565

QY 280 PTAICSFGRINQKVSNGENHVMQVNTINGTPDPDGTGVAP---LGT---PDFSGKLF 333

DB 566 GDPIDAM---IDQTV---NNQVNRSLTALQVLP--TAADTEASSHRLGTGVVPALQAAET 617

QY 334 GVLSDRDHNDACR-----SHDAVIATNSAKFTPKLGAIQIGTWEEDDVHINOPTKTPVG 388

DB 618 GASSNASDKNLTETRCVLNHHSTQETAIGNFPFSRAGLSI-----ITMPT---TG 664

QY 389 LFENEGENQWTLPNYSGALTINMGLAPPVAPTFPCEQILFPRSHIPLKGGVADPVIDCLL 448

DB 665 QONTDGYVNDI-DLMGVAQLRRKC-----ELFTYKRFDAEFVFAKPNGE-LV 712

QY 449 PQEITQHLV-QESAP---SQSDVALIRFTNPDTRVLFEAKLHRSGYITVANTGSRPIVV 504

DB 713 PQ-LIQYMYVPPGAPKPTSRDSFAQQTATNPS-----VFVKMTDPPPAQVSV-----PFMS 761

QY 505 PANGYFRFDTWVNPQYSLAP 524

DB 762 PASAY-----QW---FYDGYF 774

RESULT 11

POLG CX16G STANDARD; PRT; 2193 AA.

AC Q65900;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein

RN SEQUENCE FROM N.A.
 RP MEDLINE=87010550; PubMed=3020156;
 RA Hughes P.J., Evans D.M.A., Minor P.D., Schild G.C., Almond J.W.,
 RA Stanway G.;
 RT "The nucleotide sequence of a type 3 poliovirus isolated during a
 RT recent outbreak of poliomyelitis in Finland.";
 RL J. Gen. Virol. 67:2093-2102(1986).
 CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
 CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
 CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-I-Gly bond in the
 CC poliovirus polyprotein. In other picornavirus reactions Glu may be
 CC substituted for Gln, and Ser or Thr for Gly.
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC [RNA] (N).
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X04468; CAA28155.1; -;
 DR PIR; A27245; GNNY27.
 DR HSSP; Q84790; LPVC.
 DR MEROPS; C03.001; -;
 DR MEROPS; C03.020; -;
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR004004; Calci_pol_hel.
 DR InterPro; IPR009003; Cys_ser_trypsin.
 DR InterPro; IPR000199; Pept_3C_picorn.
 DR InterPro; IPR000081; Peptidase_C3.
 DR InterPro; IPR003138; Pico_P1A.
 DR InterPro; IPR002527; Pico_P2B.
 DR InterPro; IPR001676; Rhv.
 DR InterPro; IPR006005; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_P5vir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00548; Cys-protease-3C; 1.
 DR Pfam; PF02226; Pico_P1A; 1.
 DR Pfam; PF00947; Pico_P2A; 1.
 DR Pfam; PF01552; Pico_P2B; 1.
 DR Pfam; PF00073; rhv_3.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICVIRUSNS.
 DR ProDom; PD001125; Cys_protease_3C; 1.
 DR ProDom; PD001306; Pico_P2A; 1.
 DR ProDom; PD001274; Pico_P2B; 1.
 DR SMART; SM00382; AAA; 1.
 KW Polyprotein; Coat protein; Core protein; Transferrase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
 KW Lipoprotein.
 FT CHAIN 2 69 COAT PROTEIN VP4.
 FT CHAIN 70 340 COAT PROTEIN VP2.
 FT CHAIN 341 578 COAT PROTEIN VP1.
 FT CHAIN 579 878 COAT PROTEIN VP3.
 FT CHAIN 879 1027 CORE PROTEIN P2A.
 FT CHAIN 1028 1124 CORE PROTEIN P2B.
 FT CHAIN 1125 1453 CORE PROTEIN P2C.
 FT CHAIN 1454 1540 CORE PROTEIN P3A.
 FT CHAIN 1541 1562 GENOME-LINKED PROTEIN VPg.
 FT CHAIN 1563 1745 PICORNAIN 3C.
 FT CHAIN 1746 2206 RNA-DIRECTED RNA POLYMERASE P3D.

FT LIPID 2 2 N-myristoyl glycine (in host) (By
 FT similarity).
 FT ACT SITE 1709 1709 PROTEASE (POTENTIAL).
 FT ACT SITE 1723 1723 PROTEASE (POTENTIAL).
 SQ SEQUENCE 2206 AA; 245731 MW; F226AD85403C37BA CRC64;
 Query Match 4.6%; Score 130; DB 1; Length 2206;
 Best Local Similarity 18.7%; Pred. No. 0.23;
 Matches 114; Conservative 71; Mismatches 178; Indels 248; Gaps 29;
 QY 11 SNDGAAGLVLP-----EANNETMALEBVA-----GASIAAPLT----- 42
 DB 270 TNSATLVLPYNALSIDSMWKNWNGIALLPLDFAQDSVSRIPITVTIAPMCSEFN 329
 QY 43 GQNNIIDPWIR-LNFVQAPNG-----EFTVSPR-NSPGEV----- 75
 DB 330 GLRNVTA PKLGLPVLNTPGSGNOYLTSNQHSPCAIPEDFVTPPIDIPGEVKNVMELAEI 389
 QY 76 -----LNLLELGPENLPYLH-----LSRMVNG 98
 DB 390 DTMIPLNLENTKNTMDMYRVLSDSANLSGPICLSLSPAADPRLSHTMTLGEVLNYTH 449
 QY 99 YAGGVEVQVLLAGNAFTAGKLVFAAVPPHFPLENISPGQITWPHVIDVRLTLEPVLPL 158
 DB 450 WAGSLKFTFLFCGSMATGKLLVAAPG-AQPPTSREKAMLGTHVHDLGLQSSCTMVV 508
 QY 159 PDVRNFFHYNQNE-----PRMLVAMLYTLRSNGSGDDVFTVSCRVLTRPSP 208
 DB 509 PWISNVTYRQTQDSFTGGYISMFYQTRIVVPLSTPKAMDMLG---FVSACN----- 558
 QY 209 DFDFNYLVPPTLESKTQFTLPILTIGBLTNSRFPVIDELTYSNESLVV-----OPQNG 264
 DB 559 DFSVLLRDTT-----HISQAAMPQGVDDLITEVAQNALALS LKPKQSN 602
 QY 265 RCALDGBELQGTQLLPTALCSFRGRINOKVSGENHVMQVNTNINGTPDPDGVPAIG 324
 DB 603 -----DTRKSGPAH--SKREVTLTAVETGATN-----PL- 631
 QY 325 TPDFSGKLGVLSQLRDHD-----NACRSHDAVI-----ATNSAKETPKLGAIQ 367
 DB 632 VPSDVTQTRHVIQQRSESTIESFPARGAC---VALIEVDNEQPATNVQKL----- 680
 QY 368 IGTWE---EDDVHINQPTKFTPVGLFENEGNQWTLPNVSGALTNMGLA----- 414
 DB 681 FATWRTYTKDTVQLRRKLEFFTYSRFDME-FTFVVVTANFTNS---NNGHALNQVYQIMVI 736
 QY 415 PVAPTFFGEQILFFRSHIPLKGGVADPVIDCLLPQEWIQLHYQES-----APS 463
 DB 737 PPGAPT-----PSWDDYTWQTSNPSIIFYTGAAPA 768
 QY 464 QSDVALIRFTN 474
 DB 769 RISVFPVGLAN 779
 ID POLG HE701 STANDARD; PRT; 2194 AA.
 AC P32537;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2
 DE (P1B); Coat protein VP3 (P1C); Coat protein P2C; Core protein P3A; Genome-
 DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-
 DE linked protein VPg (P3B); Picornain 3C (IC 3.4.22.28) (Protease 3C)
 DE (P3C); RNA-directed RNA polymerase P3D (IC 3.4.22.28) (Protease 3C)].
 OS Human enterovirus 70 (strain J670/71) (EV 70).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Enterovirus.
 OX NCBI_TaxId=31915;
 RN [1]
 RP SEQUENCE FROM N.A.

Db 459 QLDAAKKSSYTLTVSAIDQANVGSRKQSSAKVKISVQDMNDNDPIFENVNKVISINENN- 517
Qy 265 RCALDGELOQTTLQTLPTAICSPGRINQKVSCHVHMNQVTNINGTFFDPTGDPVAPLG 324
Db 518 -----LAGSFVWKLTA-----KDRDSENSYISYSIANLNAVPE----- 552
Qy 325 TPDFSGLFGVLSQRDHDNACRSHDAVIATNS-----AKFTPKLGAIOIGTWEEDDVHI 378
Db 553 IDHFGS-IVKTTSLLDFFETMKNYELIIRASDWGLPYRRQTEIKLSIV-----VKDIND 605
Qy 379 NOP-----TKFTPVG 388
Db 606 NRPQFERNVCYGVTKSAPWG 626

Search completed: June 1, 2004, 13:47:21
Job time : 7.46309 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:35:27 ; Search time 31.1485 Seconds

(without alignments)
5419.275 Million cell updates/sec

Title: US-09-926-799-8

Perfect score: 2843

Sequence: 1 MKMASNDAAPSNDGAAGLVP.....VNQFYS LAPMTGTGNGRRRVQ 535

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL.25.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phase.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2843	100.0	535	12 Q9QMK6	Q9qmk6 chitta viru
2	2820	99.2	535	12 Q91I15	Q91i15 human calic
3	2814	99.0	535	12 Q917X7	Q917x7 norwalk-lik
4	2796	98.3	535	12 Q9V784	Q9v784 norwalk vir
5	2784	97.9	535	12 Q915C6	Q915c6 human calic
6	2758	97.0	535	12 Q91V39	Q91v39 human calic
7	2784	87.3	535	12 Q916E4	Q916e4 human calic
8	2477	87.1	535	12 Q917X4	Q917x4 norwalk-lik
9	2477	87.1	535	12 Q916B6	Q916b6 human calic
10	2476	87.1	535	12 Q917Y3	Q917y3 norwalk-lik
11	2474	87.0	535	12 Q68104	Q68104 hawaii cali
12	2468	86.8	535	12 Q917Y0	Q917y0 norwalk-lik
13	2466	86.7	535	12 Q917Y9	Q917y9 norwalk-lik
14	2452	86.2	535	12 Q91V51	Q91v51 human calic
15	2307.5	81.2	548	12 Q915C7	Q915c7 human calic
16	2300.5	80.9	542	12 Q913B6	Q913b6 human calic

17	2298.5	80.8	542	12 Q66293	Q66293 calicivirus
18	2296.5	80.8	542	12 Q12305	Q12305 snow mounta
19	2296.5	80.8	542	12 Q80RD6	Q80rd6 snow mounta
20	2289.5	80.5	542	12 Q96877	Q96877 snow mounta
21	2258.5	79.4	540	12 Q917X1	Q917x1 norwalk-lik
22	2255.5	79.3	540	12 Q917W8	Q917w8 norwalk-lik
23	2254.5	79.3	538	12 Q911E5	Q911e5 human calic
24	2234.5	78.6	540	12 Q91V50	Q91v50 human calic
25	2123.5	74.7	548	12 Q66296	Q66296 calicivirus
26	2123.5	74.7	548	12 Q8V768	Q8v768 norwalk vir
27	2121.5	74.6	548	12 Q8V0P5	Q8v0p5 human calic
28	2121.5	74.6	548	12 Q88291	Q88291 small round
29	2116.5	74.4	548	12 Q918A1	Q918a1 norwalk-lik
30	2116.5	74.4	548	12 Q8V0P4	Q8v0p4 human calic
31	2111.5	74.3	548	12 Q917Z5	Q917z5 norwalk-lik
32	2110.5	74.2	548	12 Q8V771	Q8v771 norwalk vir
33	2108.5	74.2	548	12 Q917Z2	Q917z2 norwalk-lik
34	2108.5	74.2	548	12 Q8BC87	Q8bc87 human calic
35	2104.5	74.0	548	12 Q83880	Q83880 norwalk vir
36	2104.5	74.0	548	12 Q91V40	Q91v40 human calic
37	2103.5	74.0	548	12 Q916E5	Q916e5 human calic
38	2101.5	73.9	548	12 Q917Z8	Q917z8 norwalk-lik
39	2099.5	73.8	548	12 Q68291	Q68291 human calic
40	2095.5	73.7	548	12 Q80FL1	Q80fl1 human calic
41	2094.5	73.7	548	12 Q8BC90	Q8bc90 human calic
42	2093.5	73.6	548	12 Q918A4	Q918a4 norwalk-lik
43	2092.5	73.6	548	12 Q80FK4	Q80fk4 human calic
44	2090.5	73.5	548	12 Q80FK9	Q80fk9 human calic
45	2090.5	73.5	548	12 Q80FK1	Q80fk1 human calic

ALIGNMENTS

RESULT 1

Q9QMK6 PRELIMINARY; PRT; 535 AA.

AC Q9QMK6; DT 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Capsid protein.

OS Chittia virus.

OC Viruses; serRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC Norovirus.

OX NCBI_TaxID=104901;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Aichi76-96;

RX MEDLINE=20471382; PubMed=11021399;

RA Kobayashi S., Sakae K., Suzuki Y., Ishiko H., Kamata K., Suzuki K.,

RA Natori K., Miyamura T., Takeda N.;

RT "Expression of recombinant capsid proteins of chitta virus, a

RT genogroup II Norwalk virus, and development of an ELISA to detect the

RT viral antigen.";

RL Microbiol. Immunol. 44:687-693 (2000).

DR EMBL; AB032758; BAA84716.1; "

DR InterPro; IPR004005; Calici_coat.

DR InterPro; IPR008975; Viral_Cap_coat.

PF Pfam; PF00915; Calici_coat; 1.

SQ SEQUENCE 535 AA; 58504 MW; 5CCEBC91BFC7B8B CRC64;

Query Match 100.0%; Score 2843; DB 12; Length 535;

Best Local Similarity 100.0%; Pred. No. 1.1e-220;

Matches 535; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY. 1 MKMASNDAAPSNDGAAGLVPPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60

Db 1 MKMASNDAAPSNDGAAGLVPPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60

QY 61 NGEFTVSPRNSPGEVLNLELGPENLPYLAHLSRMVYAGGVQVLLAGNAFTAGKLV 120

Db 61 NGEFTVSPRNSPGEVLNLELGPENLPYLAHLSRMVYAGGVQVLLAGNAFTAGKLV 120

QY 121 FAAPPHPLENISPGQITMFPVHVIDRTLEPVLPLPDPVNRNFFHYNQNEPRMRLVA 180
 DB 121 FAAPPHPLENISPGQITMFPVHVIDRTLEPVLPLPDPVNRNFFHYNQNEPRMRLVA 180
 QY 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDFNYLVPPPTLESKTKPFTLPILTIGELTNS 240
 DB 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDFNYLVPPPTLESKTKPFTLPILTIGELTNS 240
 QY 241 RFPVPIDELYSNLSVQPNQGRCALDGELOQTQLPTAICSFRCGRINQKVSNGHV 300
 DB 241 RFPVPIDELYSNLSVQPNQGRCALDGELOQTQLPTAICSFRCGRINQKVSNGHV 300
 QY 301 WNMQVNINGTDFPTDGDVPAPLGTPDFSGKLFVLSQRDHDNACRSHDAVIATNSAKFT 360
 DB 301 WNMQVNINGTDFPTDGDVPAPLGTPDFSGKLFVLSQRDHDNACRSHDAVIATNSAKFT 360
 QY 361 PKLGAIQIGTWEDDDVHINQPTKFTPVGLFENEGFNQWTLPNYSGALTINMGLAPPVAPT 420
 DB 361 PKLGAIQIGTWEDDDVHINQPTKFTPVGLFENEGFNQWTLPNYSGALTINMGLAPPVAPT 420
 QY 421 FPGQILFRSHITPLKGGVADPVIDCLLPQEWIQLHYQESAPSDVALIRFTNPTDGRV 480
 DB 421 FPGQILFRSHITPLKGGVADPVIDCLLPQEWIQLHYQESAPSDVALIRFTNPTDGRV 480
 QY 481 LFEAKLHRSYIITVANTGSRPIVVPANGYFRFDTWVNFYSLAPMGTGNGRRVQ 535
 DB 481 LFEAKLHRSYIITVANTGSRPIVVPANGYFRFDTWVNFYSLAPMGTGNGRRVQ 535
 RESULT 2
 Q91115 PRELIMINARY; PRT; 535 AA.
 AC Q91115
 DT 01-DEC-2001 (TremBrel. 19, Created)
 DT 01-DEC-2001 (TremBrel. 19, Last sequence update)
 DT 01-OCT-2003 (TremBrel. 25, Last annotation update)
 DE Capsid protein.
 OS Human calicivirus NLV/Schwerin 003/00/DE.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus
 OC NCBI_taxid=168852;
 RN SEQUENCE FROM N.A.
 RA Kuenkel U., Schreier E.;
 RT "Molecular epidemiology of outbreaks of gastroenteritis associated
 with Norwalk-like viruses in Germany."
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF397905; AAK92128.1; --
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 DR SEQUENCE 535 AA; 58520 MW; 2E97BBD3F3D5FCC1 CRC64;
 SQ
 Query Match 99.2%; Score 2820; DB 12; Length 535;
 Best Local Similarity 99.1%; Pred. No. 7.7e-219;
 Matches 530; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MKWASDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
 DB 1 MKWASDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
 QY 61 NGEFTVSPRNSGCVLLNLELGPENLPYLAHLRMYNGVAGVEVQVLLAGNAFTAGKLV 120
 DB 61 NGEFTVSPRNSGCVLLNLELGPENLPYLAHLRMYNGVAGVEVQVLLAGNAFTAGKLV 120
 QY 121 FAAPPHPLENISPGQITMFPVHVIDRTLEPVLPLPDPVNRNFFHYNQNEPRMRLVA 180
 DB 121 FAAPPHPLENISPGQITMFPVHVIDRTLEPVLPLPDPVNRNFFHYNQNEPRMRLVA 180
 QY 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDFNYLVPPPTLESKTKPFTLPILTIGELTNS 240
 DB 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDFNYLVPPPTLESKTKPFTLPILTIGELTNS 240

QY 241 RFPVPIDELYSNLSVQPNQGRCALDGELOQTQLPTAICSFRCGRINQKVSNGHV 300
 DB 241 RFPVPIDELYSNLSVQPNQGRCALDGELOQTQLPTAICSFRCGRINQKVSNGHV 300
 QY 301 WNMQVNINGTDFPTDGDVPAPLGTPDFSGKLFVLSQRDHDNACRSHDAVIATNSAKFT 360
 DB 301 WNMQVNINGTDFPTDGDVPAPLGTPDFSGKLFVLSQRDHDNACRSHDAVIATNSAKFT 360
 QY 361 PKLGAIQIGTWEDDDVHINQPTKFTPVGLFENEGFNQWTLPNYSGALTINMGLAPPVAPT 420
 DB 361 PKLGAIQIGTWEDDDVHINQPTKFTPVGLFENEGFNQWTLPNYSGALTINMGLAPPVAPT 420
 QY 421 FPGQILFRSHITPLKGGVADPVIDCLLPQEWIQLHYQESAPSDVALIRFTNPTDGRV 480
 DB 421 FPGQILFRSHITPLKGGVADPVIDCLLPQEWIQLHYQESAPSDVALIRFTNPTDGRV 480
 QY 481 LFEAKLHRSYIITVANTGSRPIVVPANGYFRFDTWVNFYSLAPMGTGNGRRVQ 535
 DB 481 LFEAKLHRSYIITVANTGSRPIVVPANGYFRFDTWVNFYSLAPMGTGNGRRVQ 535
 RESULT 3
 Q917X7 PRELIMINARY; PRT; 535 AA.
 AC Q917X7
 DT 01-DEC-2001 (TremBrel. 19, Created)
 DT 01-DEC-2001 (TremBrel. 19, Last sequence update)
 DT 01-OCT-2003 (TremBrel. 25, Last annotation update)
 DE Capsid protein.
 GN ORF2.
 OS Norwalk-like virus NLV/Honolulu/314/1994/US.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus
 OC NCBI_taxid=171852;
 RN SEQUENCE FROM N.A.
 RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
 RT "A one-tube method of reverse transcription-PCR to efficiently amplify
 a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
 of small round-structured viruses (Norwalk-like viruses).";
 RL J. Clin. Microbiol. 35:570-577(1997).
 RN [2]
 RN SEQUENCE OF 100-192 FROM N.A.
 RP STRAIN=Hu/NLV/Honolulu/314/1994/US;
 RC MEDLINE=98071277; PubMed=9407386;
 RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
 RA Seto Y., Monroe S.S., Glass R.I.;
 RT "Correlation of patient immune responses with genetically
 characterized small round-structured viruses involved in outbreaks of
 nonbacterial acute gastroenteritis in the United States, 1990 to
 1995."; J. Infect. Dis. 181:S336-S348(2000).
 RL [4]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Hu/NLV/Honolulu/314/1994/US;
 RC MEDLINE=20266071; PubMed=10804147;
 RA Ando T., Noel J.S., Fankhauser R.L.;
 RT "Genetic classification of 'Norwalk-like viruses.'";
 RL J. Infect. Dis. 181:S336-S348(2000).
 RN [4]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Hu/NLV/Honolulu/314/1994/US;
 RC Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF414420; AAL13013.1; --
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 DR SEQUENCE 535 AA; 58519 MW; C3B7A70CF3617F71 CRC64;
 SQ

Query Match 99.0%; Score 2814; DB 12; Length 535;
Best Local Similarity 98.9%; Pred. No. 2.3e-218;
Matches 529; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
DB EMBL; AB067536; BAB84137.1; -
DR EMBL; AB039775; BAC11816.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 535 AA; 58615 MW; 6B0B2D3162436787 CRC64;

QY 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMVYAGGVEVQVLLAGNAFTAGKLV 120
DB 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMVYAGGVEVQVLLAGNAFTAGKLV 120
QY 121 FFAVPPHPLENIISQGITMFPVHVIDVRLTLEPVLLPDPVNNFFHYNQNEPRMRLVA 180
DB 121 FFAVPPHPLENIISQGITMFPVHVIDVRLTLEPVLLPDPVNNFFHYNQNEPRMRLVA 180
QY 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNVLPPTLESKTKPFTLPILTIGELTNS 240
DB 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNVLPPTLESKTKPFTLPILTIGELTNS 240
QY 241 RPPVPIDELYTSPNESLVVQNGRCALDGELOTTQLLPTAICSFGRINOKVSGENHV 300
DB 241 RPPVPIDELYTSPNESLVVQNGRCALDGELOTTQLLPTAICSFGRINOKVSGENHV 300
QY 301 WNMQVNTINGTPDPTEDVPAPLGTDFSGKLFVLSQRDHDNACRSHDAVIATNSAKFT 360
DB 301 WNMQVNTINGTPDPTEDVPAPLGTDFSGKLFVLSQRDHDNACRSHDAVIATNSAKFT 360
QY 361 PKLGAIGTWEEDDVHINQKTPVGLFENEGFNQWTLPNYSGALTINMGLAPPVAPT 420
DB 361 PKLGAIGTWEEDDVHINQKTPVGLFENEGFNQWTLPNYSGALTINMGLAPPVAPT 420
QY 421 PFGEQILFFRSHIPLKGGVADPVIDCLLPQEWIQLHYQESAPSDVALIRFNTDTRGV 480
DB 421 PFGEQILFFRSHIPLKGGVADPVIDCLLPQEWIQLHYQESAPSDVALIRFNTDTRGV 480
QY 481 LFEAKLHRSYITVANTGSRPIVWPANGYFRFDTWVNFYSLAPMGTGNGRRRVQ 535
DB 481 LFEAKLHRSYITVANTGSRPIVWPANGYFRFDTWVNFYSLAPMGTGNGRRRVQ 535

RESULT 4

Q8V784 PRELIMINARY; PRT; 535 AA.
AC Q8V784;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE ORF2 protein (Capsid).
OS Norwalk virus, and
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=11983, 95340;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Norwalk virus; STRAIN=UIGII;
RA Kojima S., Kageyama T., Fukushima S., Hoshino F.B., Shinohara M.,
RA Uchida K., Natori K., Takeda N., Katayama K.;
RT "genogroup-specific primers for detect Norwalk like virus.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Norwalk-like virus; STRAIN=Saitama UI;
RA Katayama K.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=Norwalk-like virus; STRAIN=Saitama UI;
RX MEDLINE=22192455; PubMed=12202225;
RA Katayama K., Shirato-Horikoshi H., Kojima S., Kageyama T., Oka T.,
RA Hoshino F.B., Fukushima S., Shinohara M., Uchida K., Suzuki Y.,
RA Gojobori T., Takeda N.;

RT "Phylogenetic Analysis of the Complete Genome of 18 Norwalk-like
RT Viruses.";
RT Virology 299:225-239 (2002).
DR EMBL; AB067536; BAB84137.1; -
DR EMBL; AB039775; BAC11816.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 535 AA; 58615 MW; 6B0B2D3162436787 CRC64;

Query Match 98.3%; Score 2796; DB 12; Length 535;
Best Local Similarity 98.3%; Pred. No. 6.7e-217;
Matches 526; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
DB 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
QY 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMVYAGGVEVQVLLAGNAFTAGKLV 120
DB 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMVYAGGVEVQVLLAGNAFTAGKLV 120
QY 121 FFAVPPHPLENIISQGITMFPVHVIDVRLTLEPVLLPDPVNNFFHYNQNEPRMRLVA 180
DB 121 FFAVPPHPLENIISQGITMFPVHVIDVRLTLEPVLLPDPVNNFFHYNQNEPRMRLVA 180
QY 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNVLPPTLESKTKPFTLPILTIGELTNS 240
DB 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNVLPPTLESKTKPFTLPILTIGELTNS 240
QY 241 RPPVPIDELYTSPNESLVVQNGRCALDGELOTTQLLPTAICSFGRINOKVSGENHV 300
DB 241 RPPVPIDELYTSPNESLVVQNGRCALDGELOTTQLLPTAICSFGRINOKVSGENHV 300
QY 301 WNMQVNTINGTPDPTEDVPAPLGTDFSGKLFVLSQRDHDNACRSHDAVIATNSAKFT 360
DB 301 WNMQVNTINGTPDPTEDVPAPLGTDFSGKLFVLSQRDHDNACRSHDAVIATNSAKFT 360
QY 361 PKLGAIGTWEEDDVHINQKTPVGLFENEGFNQWTLPNYSGALTINMGLAPPVAPT 420
DB 361 PKLGAIGTWEEDDVHINQKTPVGLFENEGFNQWTLPNYSGALTINMGLAPPVAPT 420
QY 421 PFGEQILFFRSHIPLKGGVADPVIDCLLPQEWIQLHYQESAPSDVALIRFNTDTRGV 480
DB 421 PFGEQILFFRSHIPLKGGVADPVIDCLLPQEWIQLHYQESAPSDVALIRFNTDTRGV 480
QY 481 LFEAKLHRSYITVANTGSRPIVWPANGYFRFDTWVNFYSLAPMGTGNGRRRVQ 535
DB 481 LFEAKLHRSYITVANTGSRPIVWPANGYFRFDTWVNFYSLAPMGTGNGRRRVQ 535

RESULT 5

Q915C6 PRELIMINARY; PRT; 535 AA.
ID Q915C6
AC Q915C6;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/Pirna/110/00/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=173930;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/Pirna/110/00/DE;
RA Kuenkel U., Hoehne M., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated
RT with Norwalk-like viruses in Germany.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF427119; AAL18875.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.

```
DR Pfam; PF00915; Calici_coat; 1. D839C359F86DC1F9 CRC64;
SQ SEQUENCE 535 AA; 58505 MW; 29B513DFD2EDCFD2 CRC64;

Query Match          97.9%; Score 2784; DB 12; Length 535;
Best Local Similarity 97.4%; Pred. No. 6.2e-216;
Matches 521; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY 1 MKWASNDAAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
DB 1 MKWASNDAAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
QY 61 NGEFTVSPRNSGCVLLNLELGPENLPYLAHLSRMVNGYAGGVEQVLLAGNAFTAGKLV 120
DB 61 NGEFTVSPRNSGCVLLNLELGPENLPYLAHLSRMVNGYAGGVEQVLLAGNAFTAGKLV 120
QY 121 FFAVPPHFFLENISPGQITMFPFHVIDVLTLEPVLLPDPVRNFFHYNQNEPRMRLVA 180
DB 121 FFAVPPHFFLENISPGQITMFPFHVIDVLTLEPVLLPDPVRNFFHYNQNEPRMRLVA 180
QY 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNVLPVPTLESKTPFTLPILTIGELTNS 240
DB 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNVLPVPTLESKTPFTLPILTIGELTNS 240
QY 241 RFPVPIDELYTSNESLVVQPNQRCALDGLQGTTLQTLPTAICSFGRINQKVS GENHV 300
DB 241 RFPVPIDELYTSNESLVVQPNQRCALDGLQGTTLQTLPTAICSFGRINQKVS GENHV 300
QY 301 WNNQVTNINGTTPDPTGDPVAPLGTDFSGKLFVLSQRDHDNACRSHDAVIATNSAKFT 360
DB 301 WNNQVTNINGTTPDPTGDPVAPLGTDFSGKLFVLSQRDHDNACRSHDAVIATNSAKFT 360
QY 361 PKLGAIGTWEEDDVHINQPTKFTPVGLFENEGFNQWTLPNYSGALTLMGLAPPVAPT 420
DB 361 PKLGAIGTWEEDDVHINQPTKFTPVGLFENEGFNQWTLPNYSGALTLMGLAPPVAPT 420
QY 421 FPGSQILFFRSHIPLKGVADPVIDCLLPQEWIQLYQESAPSQSDVALIRFTNPTGRV 480
DB 421 FPGSQILFFRSHIPLKGVADPVIDCLLPQEWIQLYQESAPSQSDVALIRFTNPTGRV 480
QY 481 LFEAKLHRSYITVANTGSRPIVVPANGYFRFDTWVNFYSLAPMGTGNGRRVQ 535
DB 481 LFEAKLHRSYITVANTGSRPIVVPANGYFRFDSVWVNFYSLAPMGTGNGRRVQ 535

RESULT 6
Q91V39 PRELIMINARY; PRT; 535 AA.
ID Q91V39
AC Q91V39;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/MLV/Wortley/90/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=122922;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/MLV/Wortley/90/UK;
RX MEDLINE=20404883; PubMed=10949950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236 (2000).
DR EMBL; AJ277618; CAB89099.1; -
DR GO; GO:004601; F:peroxidase activity; IEA.
DR GO; GO:006979; P:response to oxidative stress; IEA.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR020206; Peroxidase.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR PROSITE; PS00435; PEROXIDASE 1; 1.
SQ SEQUENCE 535 AA; 58487 MW; 29B513DFD2EDCFD2 CRC64;
```

```
Query Match          97.0%; Score 2758; DB 12; Length 535;
Best Local Similarity 96.8%; Pred. No. 7.7e-214;
Matches 518; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKWASNDAAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
DB 1 MKWASNDAAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
QY 61 NGEFTVSPRNSGCVLLNLELGPENLPYLAHLSRMVNGYAGGVEQVLLAGNAFTAGKLV 120
DB 61 NGEFTVSPRNSGCVLLNLELGPENLPYLAHLSRMVNGYAGGVEQVLLAGNAFTAGKLV 120
QY 121 FFAVPPHFFLENISPGQITMFPFHVIDVLTLEPVLLPDPVRNFFHYNQNEPRMRLVA 180
DB 121 FFAVPPHFFLENISPGQITMFPFHVIDVLTLEPVLLPDPVRNFFHYNQNEPRMRLVA 180
QY 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNVLPVPTLESKTPFTLPILTIGELTNS 240
DB 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNVLPVPTLESKTPFTLPILTIGELTNS 240
QY 241 RFPVPIDELYTSNESLVVQPNQRCALDGLQGTTLQTLPTAICSFGRINQKVS GENHV 300
DB 241 RFPVPIDELYTSNESLVVQPNQRCALDGLQGTTLQTLPTAICSFGRINQKVS GENHV 300
QY 301 WNNQVTNINGTTPDPTGDPVAPLGTDFSGKLFVLSQRDHDNACRSHDAVIATNSAKFT 360
DB 301 WNNQVTNINGTTPDPTGDPVAPLGTDFSGKLFVLSQRDHDNACRSHDAVIATNSAKFT 360
QY 361 PKLGAIGTWEEDDVHINQPTKFTPVGLFENEGFNQWTLPNYSGALTLMGLAPPVAPT 420
DB 361 PKLGAIGTWEEDDVHINQPTKFTPVGLFENEGFNQWTLPNYSGALTLMGLAPPVAPT 420
QY 421 FPGSQILFFRSHIPLKGVADPVIDCLLPQEWIQLYQESAPSQSDVALIRFTNPTGRV 480
DB 421 FPGSQILFFRSHIPLKGVADPVIDCLLPQEWIQLYQESAPSQSDVALIRFTNPTGRV 480
QY 481 LFEAKLHRSYITVANTGSRPIVVPANGYFRFDTWVNFYSLAPMGTGNGRRVQ 535
DB 481 LFEAKLHRSYITVANTGSRPIVVPANGYFRFDSVWVNFYSLAPMGTGNGRRVQ 535

RESULT 7
Q916E4 PRELIMINARY; PRT; 535 AA.
ID Q916E4
AC Q916E4;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/Wiesbaden 294/01/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=173922;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/Wiesbaden 294/01/DE;
RA Kuenkel U., Hoehne M., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated
RT with Norwalk-like viruses in Germany.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
EMBL; AP425769; AAL18862.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 535 AA; 58795 MW; 987237124A0605B2 CRC64;

Query Match          87.3%; Score 2481; DB 12; Length 535;
Best Local Similarity 85.0%; Pred. No. 1.7e-191;
Matches 455; Conservative 37; Mismatches 43; Indels 0; Gaps 0;

QY 1 MKWASNDAAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
```


Db 1 MKMASNDAAPSNDGAAGLVPEVNNETMALEVPVAGASIAAPLTGQNNVIDPIRMNFVQAP 60
Qy 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMVYAGVGEVQVLLAGNAFTAGKLV 120
Db 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMVYAGVGEVQVLLAGNAFTAGKLV 120
Qy 121 FAAPVPPHLENISPGQITMPPHVIDVRLTLEPVLPLPDVNRNPFHYNQNEPRMRLVA 180
Db 121 FAAPVPPHLENISPGQITMPPHVIDVRLTLEPVLPLPDVNRNPFHYNQNEPRMRLVA 180
Qy 181 MLYTLRNSGDDVFTVSCRVLTPRSPDQFNLYLPVLTLESKTPFTLPILTIGELTNS 240
Db 181 MLYTLRNSGDDVFTVSCRVLTPRSPDQFNLYLPVLTLESKTPFTLPILTIGELTNS 240
Qy 241 RFPVDELITYSPNESLVVQPNQRCALDGLGQTTQLLTAICSFGRINQKVS GENHV 300
Db 241 RFPVDELITYSPNESLVVQPNQRCALDGLGQTTQLLTAICSFGRINQKVS GENHV 300
Qy 301 WMQVNTNNGTPTDPTGDPVAPLGTDPFSGKLVGVSORDHNAACRSHDAVATNSAKPT 360
Db 301 WMQVNTNNGTPTDPTGDPVAPLGTDPFSGKLVGVSORDHNAACRSHDAVATNSAKPT 360
Qy 361 PKLGAIGTWEEDDVHINQKTPVGLFENEGFNOWTLNYSALTLNMLAPSAPL 420
Db 361 PKLGSVLGTWEDRDFINQTRFTPVGLYDTHFNQWALPNYSALTLNMLAPSAPL 420
Qy 421 FPGQILFFRSHIPLKGGVADPVIDCLLPQEWIOHLYQESAPSDVALIRFTNPDTRGV 480
Db 421 FPGQILFFRSHIPLKGGTSNGAIDCLLPQEWVQHFYQESAPSDVALIRFTNPDTRGV 480
Qy 481 LFEAKLHRSYGITVANTGSRPIVVPANGYFRPDTWVNOFYSLAPMGTGNGRRRVQ 535
Db 481 LFEAKLHRSYGITVANTGSRPIVVPANGYFRPDTWVNOFYSLAPMGTGNGRRRVQ 535

RESULT 8
Q917X4
ID Q917X4 PRELIMINARY; PRT; 535 AA.
AC Q917X4;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Port Canaveral/301/1994/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171853;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Port Canaveral/301/1994/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify a 3-kilobase region from the RNA polymerase gene to the poly(A) tail of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Port Canaveral/301/1994/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K., Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically characterized small round-structured viruses involved in outbreaks of nonbacterial acute gastroenteritis in the United States, 1990 to 1995.";
RL J. Med. Virol. 53:372-383(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Port Canaveral/301/1994/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;

"Genetic classification of 'Norwalk-like viruses.';
J. Infect. Dis. 181:S336-S348(2000).
[4]

SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Port Canaveral/301/1994/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414421; AAL13016.1; -;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 535 AA; 58738 MW; 7C29B02DFDC724C6 CRC64;

Query Match 87.1%; Score 2477; DB 12; Length 535;
Best Local Similarity 85.0%; Pred. No. 3.6e-191;
Matches 455; Conservative 36; Mismatches 44; Indels 0; Gaps 0;

Qy 1 MKMASNDAAPSNDGAAGLVPEVNNETMALEVPVAGASIAAPLTGQNNVIDPIRMNFVQAP 60
Db 1 MKMASNDAAPSNDGAAGLVPEVNNETMALEVPVAGASIAAPLTGQNNVIDPIRMNFVQAP 60
Qy 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMVYAGVGEVQVLLAGNAFTAGKLV 120
Db 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMVYAGVGEVQVLLAGNAFTAGKLV 120
Qy 121 FAAPVPPHLENISPGQITMPPHVIDVRLTLEPVLPLPDVNRNPFHYNQNEPRMRLVA 180
Db 121 FAAPVPPHLENISPGQITMPPHVIDVRLTLEPVLPLPDVNRNPFHYNQNEPRMRLVA 180
Qy 181 MLYTLRNSGDDVFTVSCRVLTPRSPDQFNLYLPVLTLESKTPFTLPILTIGELTNS 240
Db 181 MLYTLRNSGDDVFTVSCRVLTPRSPDQFNLYLPVLTLESKTPFTLPILTIGELTNS 240
Qy 241 RFPVDELITYSPNESLVVQPNQRCALDGLGQTTQLLTAICSFGRINQKVS GENHV 300
Db 241 RFPVDELITYSPNESLVVQPNQRCALDGLGQTTQLLTAICSFGRINQKVS GENHV 300
Qy 301 WMQVNTNNGTPTDPTGDPVAPLGTDPFSGKLVGVSORDHNAACRSHDAVATNSAKPT 360
Db 301 WMQVNTNNGTPTDPTGDPVAPLGTDPFSGKLVGVSORDHNAACRSHDAVATNSAKPT 360
Qy 361 PKLGAIGTWEEDDVHINQKTPVGLFENEGFNOWTLNYSALTLNMLAPSAPL 420
Db 361 PKLGSVLGTWEDRDFINQTRFTPVGLYDTHFNQWALPNYSALTLNMLAPSAPL 420
Qy 421 FPGQILFFRSHIPLKGGVADPVIDCLLPQEWIOHLYQESAPSDVALIRFTNPDTRGV 480
Db 421 FPGQILFFRSHIPLKGGTSNGAIDCLLPQEWVQHFYQESAPSDVALIRFTNPDTRGV 480
Qy 481 LFEAKLHRSYGITVANTGSRPIVVPANGYFRPDTWVNOFYSLAPMGTGNGRRRVQ 535
Db 481 LFEAKLHRSYGITVANTGSRPIVVPANGYFRPDTWVNOFYSLAPMGTGNGRRRVQ 535

RESULT 9

Q916E6
ID Q916E6 PRELIMINARY; PRT; 535 AA.
AC Q916E6;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/Dillingen 391/01/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=173920;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/Dillingen 391/01/DE;
RA Kuenkel U., Hoehne M., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Bates M.K.,
RA Seto Y., Monroe S.S., Glass R.I.,
RT "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995.";
RL J. Med. Virol. 53:372-383 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Westover/302/1994/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk'-like viruses.";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Westover/302/1994/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF414418; AAL13007.1;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 535 AA; 58684 MW; 62620468AB44EC0 CRC64;

Query Match 87.1%; Score 2477; DB 12; Length 535;
Best Local Similarity 84.9%; Pred. No. 3.6e-191;
Matches 454; Conservative 37; Mismatches 44; Indels 0; Gaps 0;
QY 1 MKMASNDAAPSNDGAAGLVPEANNEITMALEPVAGASIAAAPTGTQNNIIDPWIRNFVQAP 60
Db 1 MKMASNDAAPSNDGAAGLVPEANNEITMALEPVAGASIAAAPTGTQNNIIDPWIRNFVQAP 60
QY 61 NGEFTVSPNSGCEVLNLELGPENLPYLAHLSRMVNGYAGGVEVQVLLAGNAFTAGKLV 120
Db 61 NGEFTVSPNSGCEVLNLELGPENLPYLAHLSRMVNGYAGGVEVQVLLAGNAFTAGKLV 120
QY 121 FAAPVPHPELENISPGQITMFPFHVIIIDVRLTPVLLPLPDVVRNFFHYNQNEPRMRLVA 180
Db 121 FAAPVPHPELENISPGQITMFPFHVIIIDVRLTPVLLPLPDVVRNFFHYNQNEPRMRLVA 180
QY 181 MLYTPLRSNGSGDDVFTVSCRVLTRPSDFDFNYLVPPTLESKTKPFTLPILITIGELTNS 240
Db 181 MLYTPLRSNGSGDDVFTVSCRVLTRPSDFDFNYLVPPTLESKTKPFTLPILITIGELTNS 240
QY 241 RFPVPIDELTSPNESLVVQPNQRCALDGELOQTTLPTAICSPRGRINOKVSGENHV 300
Db 241 RFPVPIDELTSPNESLVVQPNQRCALDGELOQTTLPTAICSPRGRINOKVSGENHV 300
QY 301 WNMQVNINGTPDPTGDPAPLGTDPFSGKLFVLSQRDHNA CRSHDAVIATNSAKFT 360
Db 301 WNMQVNINGTPDPTGDPAPLGTDPFSGKLFVLSQRDHNA CRSHDAVIATNSAKFT 360
QY 361 PKLGAIQIGTWEEEDVHINQPTKFTPVGLFENEGFNQWTLNPNYSGALTNNGLAPPVAPT 420
Db 361 PKLGSVVLGTWEDRDINQPTKFTPVGLYDTHFNQWALPNYSGALTNNGLAPPVAPT 420
QY 421 FPGEQILFFRSHIPLKGGVADPVDCLLPQEWIQLYQESAPSDVALIRFTNPDTGRV 480
Db 421 FPGEQILFFRSHIPLKGGTNGAIDCLLPQEWVQHFYQESAPSDTVALIRFTNPDTGRV 480
QY 481 LFEAKLHRSGYITVANTGSRPIVVPNGYFRFDSVWVNFYSLAPMGTCNGRRRVQ 535
Db 481 LFEAKLHRSQGITVANSGRPIVVPNGYFRFDSVWVNFYSLAPMGTCNGRRRVQ 535
RESULT 10
Q917Y3 ID Q917Y3 PRELIMINARY; PRT; 535 AA.
AC Q917Y3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Westover/302/1994/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171850;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Westover/302/1994/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577 (1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Westover/302/1994/US;
RX MEDLINE=98071277; PubMed=9407386;

QY 1 MKMASNDAAPSNDGAAGLVPEANNEITMALEPVAGASIAAAPTGTQNNIIDPWIRNFVQAP 60
Db 1 MKMASNDAAPSNDGAAGLVPEANNEITMALEPVAGASIAAAPTGTQNNIIDPWIRNFVQAP 60
QY 61 NGEFTVSPNSGCEVLNLELGPENLPYLAHLSRMVNGYAGGVEVQVLLAGNAFTAGKLV 120
Db 61 NGEFTVSPNSGCEVLNLELGPENLPYLAHLSRMVNGYAGGVEVQVLLAGNAFTAGKLV 120
QY 121 FAAPVPHPELENISPGQITMFPFHVIIIDVRLTPVLLPLPDVVRNFFHYNQNEPRMRLVA 180
Db 121 FAAPVPHPELENISPGQITMFPFHVIIIDVRLTPVLLPLPDVVRNFFHYNQNEPRMRLVA 180
QY 181 MLYTPLRSNGSGDDVFTVSCRVLTRPSDFDFNYLVPPTLESKTKPFTLPILITIGELTNS 240
Db 181 MLYTPLRSNGSGDDVFTVSCRVLTRPSDFDFNYLVPPTLESKTKPFTLPILITIGELTNS 240
QY 241 RFPVPIDELTSPNESLVVQPNQRCALDGELOQTTLPTAICSPRGRINOKVSGENHV 300
Db 241 RFPVPIDELTSPNESLVVQPNQRCALDGELOQTTLPTAICSPRGRINOKVSGENHV 300
QY 301 WNMQVNINGTPDPTGDPAPLGTDPFSGKLFVLSQRDHNA CRSHDAVIATNSAKFT 360
Db 301 WNMQVNINGTPDPTGDPAPLGTDPFSGKLFVLSQRDHNA CRSHDAVIATNSAKFT 360
QY 361 PKLGAIQIGTWEEEDVHINQPTKFTPVGLFENEGFNQWTLNPNYSGALTNNGLAPPVAPT 420
Db 361 PKLGSVVLGTWEDRDINQPTKFTPVGLYDTHFNQWALPNYSGALTNNGLAPPVAPT 420
QY 421 FPGEQILFFRSHIPLKGGVADPVDCLLPQEWIQLYQESAPSDVALIRFTNPDTGRV 480
Db 421 FPGEQILFFRSHIPLKGGTNGAIDCLLPQEWVQHFYQESAPSDTVALIRFTNPDTGRV 480
QY 481 LFEAKLHRSGYITVANTGSRPIVVPNGYFRFDSVWVNFYSLAPMGTCNGRRRVQ 535
Db 481 LFEAKLHRSQGITVANSGRPIVVPNGYFRFDSVWVNFYSLAPMGTCNGRRRVQ 535

RESULT 11
Q68104 ID Q68104 PRELIMINARY; PRT; 535 AA.
AC Q68104;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Capsid protein.
OS Hawaii calicivirus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=33750;
RN [1]_TaxID=33750;
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;
RX MEDLINE=94358485; PubMed=8077710;
RA Lew J.F., Kapikian A.Z., Valdesuso J., Green K.Y.;
RT "Molecular characterization of Hawaii virus and other Norwalk-like
RT viruses; evidence for genetic polymorphism among human
RT caliciviruses."; J. Infect. Dis. 170:535-542 (1994).
RL J. Infect. Dis. 170:535-542 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;
RX MEDLINE=94358485; PubMed=8077710;
RA Lew J.F.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;
RX MEDLINE=94358485; PubMed=8077710;
RA Green K.Y., Sosnovtsev S.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;
RX MEDLINE=94358485; PubMed=8077710;
RA Pleitneva M.A., Sosnovtsev S., Green K.Y.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; U07611; AAB97768.2; -;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 535 AA; 58741 MW; D0CA252A210E4599 CRC64;

Query Match 87.0%; Score 2474; DB 12; Length 535;
Best Local Similarity 84.9%; Pred. No. 6.3e-191;
Matches 454; Conservative 40; Mismatches 41; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIDPWIRLNFVQAP 60
DB 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIDPWIRLNFVQAP 60

QY 61 NGEFTVSPNSPCEVLLNLELQPELNPYLAHLRMYNGYAGVEVQVLLAGNAFTAGKLV 120
DB 61 NGEFTVSPNSPCEVLLNLELQPELNPYLAHLRMYNGYAGVEVQVLLAGNAFTAGKLV 120

QY 121 FAAPVPHPLENISQGITMFPVHVIDVRLTLEPVLPLPDVRNPFHYNQOEPRMRLVA 180
DB 121 FAAPVPHPLENISQGITMFPVHVIDVRLTLEPVLPLPDVRNPFHYNQOEPRMRLVA 180

QY 181 MLYTPLRSNGSGDDVFTVSCRVLTRPSDFDNYLVPPTLESKTKPFTLPILTIGELTNS 240
DB 181 MLYTPLRSNGSGDDVFTVSCRVLTRPSDFDNYLVPPTLESKTKPFTLPILTIGELTNS 240

QY 241 RFPVPIDELYTSFNBSLVVQPNQRCALDGEQGTQLLPTAICSPGRINOKVSGENHV 300
DB 241 RFPVPIDELYTSFNBSLVVQPNQRCALDGEQGTQLLPTAICSPGRINOKVSGENHV 300

QY 301 WNVQVNTINGTPTDPTGDPVAPLPTDPSGKLVGLVSRDHDNACRSHDAVATNSAKPT 360
DB 301 WNVQVNTINGTPTDPTGDPVAPLPTDPSGKLVGLVSRDHDNACRSHDAVATNSAKPT 360

QY 361 PKLGAIOGTWEEDVHINOPTKFTPVGLFENEGFNQWTLPNYSGALTLNMGGLAPVAPT 420
DB 361 PKLGSVILGTWESDLDLNOPTFRFTVGLFNTDHDQWALPSGRGLTLNMLNLSVSL 420

QY 421 FPEGQILFRSHIPLKGGVADPVIDCLLPQEWIQLHYQESAPSQSVALIRTNPDTRGV 480
DB 421 FPEGQILFRSHIPLKGGVADPVIDCLLPQEWIQLHYQESAPSQSVALIRTNPDTRGV 480

QY 481 LFEAKLHRSGYITVANTGSRPIVWPANGYFRFDTWVNFQYSLAPMGTCNGRRRVQ 535
DB 481 LFEAKLHRSGYITVANTGSRPIVWPANGYFRFDTWVNFQYSLAPMGTCNGRRRVQ 535

Db 481 LFEAKLHRSGYITVANTGSRPIVWPANGYFRFDTWVNFQYSLAPMGTCNGRRRVQ 535

RESULT 12
Q917Y0 PRELIMINARY; PRT; 535 AA.
AC Q917Y0;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Richmond/283/1994/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171851;
RN [1]_TaxID=171851;
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Richmond/283/1994/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577 (1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Richmond/283/1994/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
RA Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995."; J. Med. Virol. 53:372-383 (1997).
RL J. Med. Virol. 53:372-383 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Richmond/283/1994/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk'-like viruses.";
RL J. Infect. Dis. 181:S336-S348 (2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Richmond/283/1994/US;
RA Ando T., Seto Y., Noel J.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414419; AAL13010.1; -;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 535 AA; 58742 MW; 4EDACABCB8E22F80 CRC64;

Query Match 86.8%; Score 2468; DB 12; Length 535;
Best Local Similarity 84.9%; Pred. No. 1.9e-190;
Matches 454; Conservative 36; Mismatches 45; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIDPWIRLNFVQAP 60
DB 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIDPWIRLNFVQAP 60

QY 61 NGEFTVSPNSPCEVLLNLELQPELNPYLAHLRMYNGYAGVEVQVLLAGNAFTAGKLV 120
DB 61 NGEFTVSPNSPCEVLLNLELQPELNPYLAHLRMYNGYAGVEVQVLLAGNAFTAGKLV 120

QY 121 FAAPVPHPLENISQGITMFPVHVIDVRLTLEPVLPLPDVRNPFHYNQOEPRMRLVA 180
DB 121 FAAPVPHPLENISQGITMFPVHVIDVRLTLEPVLPLPDVRNPFHYNQOEPRMRLVA 180

QY 181 MLYTPLRSNGSGDDVFTVSCRVLTRPSDFDNYLVPPTLESKTKPFTLPILTIGELTNS 240
DB 181 MLYTPLRSNGSGDDVFTVSCRVLTRPSDFDNYLVPPTLESKTKPFTLPILTIGELTNS 240

QY 241 RPPVPIDELYTSPNESLWVQNGRCALDGELOCTTOLLPTAICSFRRINOKVSGENHV 300
DB 241 RPPVPIDELYTSPNESLWVQNGRCALDGELOCTTOLLPTAICSFRRINOKVSGENHV 300
QY 301 WNMQVNTNNGTFFDPTGDPVAPPLGTPDFSGKLFVLSQRDHDNACRSHDAVIATNSAKPT 360
DB 301 WNMQVNTNNGTFFDPTGDPVAPPLGTPDFSGKLFVLSQRDHDNACRSHDAVIATNSAKPT 360
QY 361 PKLGAIOIGTWEDDDVHINQPTKFTPVGLFENEGFNQWTLNYSGALTINMGLAPPVAPT 420
DB 361 PKLGSVVLGTWEDRDFDINQPARFETPVGLYDTHFNQWALPNYSGALTINMGLAPPVAPT 420
QY 421 FPGEOQLFRSHIPLKGGVADPVIDCLLPQEWIQHLYQESAPSDQDVALIRFTNPDTCRV 480
DB 421 FPGEOQLFRSHIPLKGGVADPVIDCLLPQEWIQHLYQESAPSDQDVALIRFTNPDTCRV 480
QY 481 LFEAKLHRSQGIITVANTGSRPIVVPANGYFRDPTWVQNYFSLAPMGTGNGRRVQ 535
DB 481 LFEAKLHRSQGIITVANTGSRPIVVPANGYFRDPTWVQNYFSLAPMGTGNGRRVQ 535
RESULT 13
Q917Y9 PRELIMINARY; PRT; 535 AA.
AC Q917Y9
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Miami/81/1986/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_taxid=171848;
RN [1]
RC STRAIN=Hu/NLV/Miami/81/1986/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RC STRAIN=Hu/NLV/Miami/81/1986/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [3]
RC STRAIN=Hu/NLV/Miami/81/1986/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414416; AAL13001.1;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 535 AA; 58745 MW; D43CDF7BD8C7225A CRC64;
Query Match 86.7%; Score 2466; DB 12; Length 535;
Best Local Similarity 84.7%; Pred. No. 2.8e-190;
Matches 453; Conservative 37; Mismatches 45; Indels 0; Gaps 0;
QY 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
DB 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
QY 61 NGEFTVSPRNSPQEVLLNLELGPENLPLAHLRSMYNGYAGGVQVLLAGNAFTAGKLV 120
DB 61 NGEFTVSPRNSPQEVLLNLELGPENLPLAHLRSMYNGYAGGVQVLLAGNAFTAGKLV 120

QY 121 FAAVPDPHLENI SPQITMFPVHVIDVRLTEPVLPLDVRNFFHYNQNEPRMLVA 180
DB 121 FAAVPDPHLENI SPQITMFPVHVIDVRLTEPVLPLDVRNFFHYNQNEPRMLVA 180
QY 181 MLYTPLRNGSGDDVFTVSCRVLTSPDPDFPNLVPTLESKTPFTLPTLTIGELTNS 240
DB 181 MLYTPLRNGSGDDVFTVSCRVLTSPDPDFPNLVPTLESKTPFTLPTLTIGELTNS 240
QY 241 RFPVPIDELYTSPNESLWVQNGRCALDGELOCTTOLLPTAICSFRRINOKVSGENHV 300
DB 241 RFPVPIDELYTSPNESLWVQNGRCALDGELOCTTOLLPTAICSFRRINOKVSGENHV 300
QY 301 WNMQVNTNNGTFFDPTGDPVAPPLGTPDFSGKLFVLSQRDHDNACRSHDAVIATNSAKPT 360
DB 301 WNMQVNTNNGTFFDPTGDPVAPPLGTPDFSGKLFVLSQRDHDNACRSHDAVIATNSAKPT 360
QY 361 PKLGAIOIGTWEDDDVHINQPTKFTPVGLFENEGFNQWTLNYSGALTINMGLAPPVAPT 420
DB 361 PKLGSVVLGTWEDRDFDINQPARFETPVGLYDTHFNQWALPNYSGALTINMGLAPPVAPT 420
QY 421 FPGEOQLFRSHIPLKGGVADPVIDCLLPQEWIQHLYQESAPSDQDVALIRFTNPDTCRV 480
DB 421 FPGEOQLFRSHIPLKGGVADPVIDCLLPQEWIQHLYQESAPSDQDVALIRFTNPDTCRV 480
QY 481 LFEAKLHRSQGIITVANTGSRPIVVPANGYFRDPTWVQNYFSLAPMGTGNGRRVQ 535
DB 481 LFEAKLHRSQGIITVANTGSRPIVVPANGYFRDPTWVQNYFSLAPMGTGNGRRVQ 535
RESULT 14
Q91V51 PRELIMINARY; PRT; 535 AA.
AC Q91V51
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/Ludwigslust/218/99/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=122910;
RN [1]
RC STRAIN=Hu/NLV/Girlington/93/UK;
RX MEDLINE=20404883; PubMed=10949950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL; AJ277606; CAB89087.1;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 535 AA; 58772 MW; 27C6CC7DBFC2A6AB CRC64;
Query Match 86.2%; Score 2452; DB 12; Length 535;
Best Local Similarity 84.3%; Pred. No. 3.7e-189;
Matches 451; Conservative 37; Mismatches 47; Indels 0; Gaps 0;
QY 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
DB 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
QY 61 NGEFTVSPRNSPQEVLLNLELGPENLPLAHLRSMYNGYAGGVQVLLAGNAFTAGKLV 120
DB 61 NGEFTVSPRNSPQEVLLNLELGPENLPLAHLRSMYNGYAGGVQVLLAGNAFTAGKLV 120
QY 121 FAAVPDPHLENI SPQITMFPVHVIDVRLTEPVLPLDVRNFFHYNQNEPRMLVA 180
DB 121 FAAVPDPHLENI SPQITMFPVHVIDVRLTEPVLPLDVRNFFHYNQNEPRMLVA 180
QY 181 MLYTPLRNGSGDDVFTVSCRVLTSPDPDFPNLVPTLESKTPFTLPTLTIGELTNS 240
DB 181 MLYTPLRNGSGDDVFTVSCRVLTSPDPDFPNLVPTLESKTPFTLPTLTIGELTNS 240

```
Db 181 MLYTPLRNSGDDVFTVSCRPTSPDFDFNYLVPPTVESKTPFTLPILTIGELNS 240
Qy 241 RPPVPIDELYSNLSVQPNQRCALDGEIQQGTTQLLPTAICSPRGRINQKVSGENHV 300
Db 241 RPPVPIDELYSNLSVQPNQRCALDGEIQQGTTQLLPTAICSPRGRINQKVSGENHV 300
Qy 301 WNMQVNTINGTPTDPTGDPVAPLGTDFSGKLFVLSQRDHNAACRSHDAVIATNSAKPT 360
Db 301 WNMQVNTANGTPTDPTEDVPALGTDFSLANIYGVTSQRNPDNTCRAHDGILATWSPKPT 360
Qy 361 PKLGAIGTWEEDVHINQKFTPVGLFENEGNQWTLPNYSGALTINMGLAPVAPT 420
Db 361 PKLGSVVLGTWEDRDVDINQKFTPVGLFENEGNQWTLPNYSGALTINMGLAPVAPT 420
Qy 421 FPEGQILFRSHPIPLKGGVADVIDCLLPQEWIQLYQESAPSDVALIRFTNPDTRGV 480
Db 421 FPEGQILFRSHPIPLKGGTSNGAIDCLLPQEWVQHFYQESAPSDVALIRFTNPDTRGV 480
Qy 481 LFEAKLHRSGYITVANTGSRPIVWPANGYFRDFTWVNOFYSLAPMGTGNRRRVQ 535
Db 481 LFEAKLHRCGFTVANTGSRPIVWPANGYFRDFTWVNOFYSLAPMGTGNRRRVQ 535
```

RESULT 15

```
Q915C7 PRELIMINARY; PRT; 548 AA.
ID Q915C7;
AC Q915C7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/Exfurt/546/00/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_Taxid=173929;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/Exfurt/546/00/DE;
RA Kuenkel U., Hoehne M., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated
RT with Norwalk-like viruses in Germany.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF427118; AAL18874.1; -
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 59874 MW; 9A1D8BEE58ADF7D CRC64;
```

```
Query Match 81.2%; Score 2307.5; DB 12; Length 548;
Best Local Similarity 77.7%; Pred. No. 1.7e-177;
Matches 426; Conservative 57; Mismatches 52; Indels 13; Gaps 3;
Qy 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
Db 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
Qy 61 NGEFTVSPRNSPCEVLNLELGPENLNPYLALSRMTNGYAGGVQVLLAGNAFTAGKLV 120
Db 61 NGEFTVSPRNSPCEVLNLELGPENLNPYLALSRMTNGYAGGVQVLLAGNAFTAGKLV 120
Qy 121 FAAPVPHFPLENISQGITWTFPHVVIDVRLTLPVLLPLDPVRNPFPHYNQONEPRMLVA 180
Db 121 FAAPVPHFPLENISQGITWTFPHVVIDVRLTLPVLLPLDPVRNPFPHYNQONEPRMLVA 180
Qy 181 MLYTPLRNSGDDVFTVSCRVLTRPSPDFDFNYLVPPTLESKTPFTLPILTIGELTNS 240
Db 181 MLYTPLRNSGDDVFTVSCRVLTRPSPDFDFNYLVPPTLESKTPFTLPILTIGELTNS 240
Qy 241 RPPVPIDELYSNLSVQPNQRCALDGEIQQGTTQLLPTAICSPRGRINQKVSGENH- 299
Db 241 RPPVLPIDVLYTNPNESAIVOCQNGRCTLDBGELQGTQLLPTGICAPRGKVTQOVQDEHRG 300
```

```
Qy 300 -VWMQVNTINGTPTDPTGDPVAPLGTDFSGKLFVLSQRDHNAACRSHDA 350
Db 301 THMNTVNTLANGTPTDPTEDVPALGTDFSGIYGVISQRNNTVTPGEGDLPANRAHEA 360
Qy 351 VIATNSAKFTPKLGAIGTWEEDVHINQKFTPVGLFE---NEGFNQWTLPNYSGAL 407
Db 351 VIATNSAKFTPKLGNIOFSTWETQDVSSGQPTKFTPVGLASVDANSHDFDQWTLPSYSGAL 420
Qy 408 TLNMGLAPVAPTFFPEQILFRSHPIPLKGGVADVIDCLLPQEWIQLYQESAPSDV 467
Db 421 TLNMGLAPVAPTFFPEQILFRSHPIPLKGGYGNDAIDCLMPQEWVQHFYQESAPSLSDV 480
Qy 468 ALIRFTNPDTRGVLPQEWIQLYQESAPSDVALIRFTNPDTRGV 527
Db 481 ALVRYNPETGRTLFEAKLHRCGFTVANTGSRPIVWPANGYFRDFTWVNOFYSLAPMGT 540
Qy 528 GNGRRRVQ 535
Db 541 GSGRRMQ 548
```

Search completed: June 1, 2004, 13:53:20
Job time : 32.1485 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time, 46.5611 Seconds
(without alignments)
3289.030 Million cell updates/sec

Title: US-09-926-799-9
Perfect score: 2863
Sequence: 1 MKMASNDAPSDGASLVP.....VNQFYSLPAPMTGNGRRRIQ 542

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2863	100.0	542	4	AAB49708 Small rou
2	2048.5	71.6	535	4	AAB49707 Small rou
3	1933	67.5	540	4	AAB49706 Small rou
4	1931	67.4	548	4	AAB49705 Small rou
5	1886	65.9	548	5	AAB49705 Small rou
6	1872	65.4	550	4	AAB49709 Small rou
7	1821.5	63.6	539	4	AAB49704 Small rou
8	1538.5	53.7	541	4	AAB49710 Small rou
9	1181	41.3	530	4	AAB49701 Small rou
10	1179	41.2	530	2	AAB50972 Norwalk v
11	1179	41.2	530	7	ADC72176 Norwalk v
12	1172.5	41.0	545	4	AAB49700 Small rou
13	1165.5	40.7	544	4	AAB49703 Small rou
14	1155.5	40.4	546	4	AAB49702 Small rou
15	1137	39.7	530	2	AAB57091 Small rou
16	336	11.7	579	2	AAB08143 RHDV caps
17	279	9.7	668	2	AAB10686 Feline ca
18	279	9.7	668	4	AAB04304 Feline ca
19	278	9.7	623	4	AAB47044 Feline ca
20	278	9.7	623	4	AAB47043 Feline ca
21	276.5	9.7	547	4	AAB50108 Feline ca
22	276.5	9.7	671	4	AAB50107 Feline ca
23	275	9.6	622	4	AAB47045 Feline ca
24	265.5	9.3	668	4	AAB67462 Amino aci
25	265.5	9.3	669	4	AAB67461 Amino aci

ALIGNMENTS

RESULT 1
AAB49708
ID AAB49708 standard; protein; 542 AA.

AC AAB49708;
XX
DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 9.
XX
KW Small round structured virus; SRSV; food poisoning.
XX
OS Small round structured virus.
XX
PN WO200079280-A1.
XX
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000MO-JP004095.
XX
PR 22-JUN-1999; 95JP-00175928.
XX
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
WPI, 2001-080848/09.
N-PSDB; AAF29149.

Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.

Claim 1; Page 59-61; 84pp; Japanese.

This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

Sequence 542 AA;

Query Match 100.0%; Score 2863; DB 4; Length 542;
Best Local Similarity 100.0%; Pred. No. 2.2e-237;

Matches 542; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MKWASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQTNIIIDPWIRTNFVOAP 60
DB	1 MKWASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQTNIIIDPWIRTNFVOAP 60
QY	61 NGEFTVSPRNSPGEILLNLLELGPDLNPNYLAHLSRMYNGYAGGVEVQVLLAGNAFTAGKIL 120
DB	61 NGEFTVSPRNSPGEILLNLLELGPDLNPNYLAHLSRMYNGYAGGVEVQVLLAGNAFTAGKIL 120
QY	121 FAAIPNFIIVDMISPAQITMLPHLIVDVRTLEPIIMTLPDVRNVFHFNNQPPQMRLLVA 180
DB	121 FAAIPNFIIVDMISPAQITMLPHLIVDVRTLEPIIMTLPDVRNVFHFNNQPPQMRLLVA 180
QY	181 MLYTPLRSNGSGDDVFTVSCRVLTRPTDPFEIYLVPPSVESKTKPFTLPILTISELTNS 240
DB	181 MLYTPLRSNGSGDDVFTVSCRVLTRPTDPFEIYLVPPSVESKTKPFTLPILTISELTNS 240
QY	241 RFPPIEQLYTAPNETNVVQCNGRCCTLDGELQGTQLLSSAVCFIQGRTVADNGDNWDQ 300
DB	241 RFPPIEQLYTAPNETNVVQCNGRCCTLDGELQGTQLLSSAVCFIQGRTVADNGDNWDQ 300
QY	301 NLLQLYTPNGASVDPDTEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGIYISITS 360
DB	301 NLLQLYTPNGASVDPDTEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGIYISITS 360
QY	361 GKFTPKIGSIGLSITTEHVHPNQSRFTPVGVAVDENTPFQWVLPHYAGSLALNTNLAP 420
DB	361 GKFTPKIGSIGLSITTEHVHPNQSRFTPVGVAVDENTPFQWVLPHYAGSLALNTNLAP 420
QY	421 AVAPTFPGEQLLFFRSRVPVQGLQGDQAFIDCLLPQEWVNHFYQEAAPSQADVALIRYV 480
DB	421 AVAPTFPGEQLLFFRSRVPVQGLQGDQAFIDCLLPQEWVNHFYQEAAPSQADVALIRYV 480
QY	481 NPDTGRTLFEAKLHRSQFIIVSHGTGAYPLVPPNGHFRFDSWVNFYSLAPMGNGRRR 540
DB	481 NPDTGRTLFEAKLHRSQFIIVSHGTGAYPLVPPNGHFRFDSWVNFYSLAPMGNGRRR 540
QY	541 IQ 542
DB	541 IQ 542
RESULT 2	
ID	AAB49707
XX	AAB49707 standard; protein; 535 AA.
AC	AAB49707;
XX	
DT	04-APR-2001 (first entry)
XX	
DE	Small round structured virus protein SEQ ID 8.
XX	
KW	Small round structured virus; SRSV; food poisoning.
XX	
OS	Small round structured virus.
XX	
PN	WO200079280-A1.
XX	
PD	28-DEC-2000.
XX	
PF	22-JUN-2000; 2000WO-JP004095.
XX	
PR	22-JUN-1999; 99JP-00175928.
XX	
PA	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX	
PI	(DENK-) DENKA SEIKEN KK.
XX	
DR	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX	
DR	WPI; 2001-080848/09.
XX	
DR	N-PSDB; AAF29148.
XX	

Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.

Claim 1; Page 57-59; 84pp; Japanese.

This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

Sequence 535 AA;

Query Match 71.6%; Score 2048.5; DB 4; Length 535;
Best Local Similarity 71.7%; Pred. No. 2.9e-167; Indels 17; Gaps 7;
Matches 392; Conservative 64; Mismatches 74;

QY 1 MKWASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQTNIIIDPWIRTNFVOAP 60
DB 1 MKWASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQTNIIIDPWIRTNFVOAP 60
QY 61 NGEFTVSPRNSPGEILLNLLELGPDLNPNYLAHLSRMYNGYAGGVEVQVLLAGNAFTAGKIL 120
DB 61 NGEFTVSPRNSPGEILLNLLELGPDLNPNYLAHLSRMYNGYAGGVEVQVLLAGNAFTAGKIL 120
QY 121 FAAIPNFIIVDMISPAQITMLPHLIVDVRTLEPIIMTLPDVRNVFHFNNQPPQMRLLVA 180
DB 121 FAAIPNFIIVDMISPAQITMLPHLIVDVRTLEPIIMTLPDVRNVFHFNNQPPQMRLLVA 180
QY 181 MLYTPLRSNGSGDDVFTVSCRVLTRPTDPFEIYLVPPSVESKTKPFTLPILTISELTNS 240
DB 181 MLYTPLRSNGSGDDVFTVSCRVLTRPTDPFEIYLVPPSVESKTKPFTLPILTISELTNS 240
QY 241 RFPPIEQLYTAPNETNVVQCNGRCCTLDGELQGTQLLSSAVCFIQGRTVADNGDNWDQ 300
DB 241 RFPPIEQLYTAPNETNVVQCNGRCCTLDGELQGTQLLSSAVCFIQGRTVADNGDNWDQ 300
QY 298 WDQNLQLTYTPNGASVDPDTEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGIYI 356
DB 301 MN---MQVTNINGTFFDGTGVPAPLGTDFSGMLYGVLTQDNVNVSTGEAKNAKGIYI 352
QY 357 STTSCKFTPKIGSIGLSITTEHVHPNQSRFTPVGVAVDENTPFQWVLPHYAGSLALN 415
DB 353 ATNSAKFTPKLGAIQIGTWEEDDVHINQPTKFTPVGLF--ENEGFNQWTLFNYSGALTLN 410
QY 416 TNLAPAVAPTFPGEQLLFFRSRVPVQGLQGDQAFIDCLLPQEWVNHFYQEAAPSQADVA 475
DB 411 MGLAPPVAPTFPGEQLLFFRSRVPVQGLQGDQAFIDCLLPQEWVNHFYQEAAPSQADVA 468
QY 476 LIRYVNPDTGRTLFEAKLHRSQFIIVSHGTGAYPLVPPNGHFRFDSWVNFYSLAPMGNG 535
DB 469 LIRFTNPDTGRTLFEAKLHRSQFIIVSHGTGAYPLVPPNGHFRFDSWVNFYSLAPMGNG 528
QY 536 NGRRIQ 542
DB 529 NGRRRVQ 535

RESULT 3
AAB49706
ID AAB49706 standard; protein; 540 AA.
XX
AC AAB49706;
XX
DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 7.
XX
KW Small round structured virus; SRSV; food poisoning.
XX

OS Small round structured virus.
 PN WO200079280-A1.
 XX 28-DEC-2000.
 PD
 XX 22-JUN-2000; 2000WO-JP004095.
 PF
 XX 22-JUN-1999; 99JP-00175928.
 PR
 XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.
 XX
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 XX WPI; 2001-080848/09.
 DR N-PSDB; AAF29147.
 XX
 XX Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX
 XX Claim 1; Page 54-57; 84pp; Japanese.
 PS
 XX This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 CC
 XX Sequence 540 AA;
 SQ
 Query Match 67.5%; Score 1933; DB 4; Length 540;
 Best Local Similarity 67.0%; Pred. No. 2.6e-157; Mismatches 94; Indels 10; Gaps 6;
 Matches 366; Conservative 76;
 QY 1 MKWASNDAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQTNIIIDPWIRTNFVQAP 60
 DB 1 MKWASNDATPSNDGAAGLVPESSNNEAMALEPVVGSALAAPVTGQTNIIIDPWIRTNFVQAP 60
 QY 61 NGEFTVSPRNSGCEILLNLELGPDLNPLYLAHLSRMVNGYAGGVQVLLAGNAFTAGKIL 120
 DB 61 NGEFTVSPRNSGCEILLNLELGPDLNPLYLAHLSRMVNGYAGGVQVLLAGNAFTAGKII 120
 QY 121 FAAPPPFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNPFVHNNQPPQMRLLVA 180
 DB 121 FAAPPPFPFVNLSPSQITMFPVHVIDVRLTLEPVLPMVDVRSSTLTFHFNQDEPKRLLVA 180
 QY 181 MLYTPLRNSGDDVFTVSCRVLTPTDPFETIYLVPPSVESKTKPFTLPILTITSELNLS 240
 DB 181 MLYTPLRNSGDDVFTVSCRVLTPTDPFETIYLVPPSVESKTKPFTLPILTITSELNLS 240
 QY 241 RFPPIEQLYTAPNETNVVQCNQGRCTLDGELQGTQTLSSAVCFLOGRT--VADNGDN 297
 DB 241 RFPPLSIDEMVTSFNESIVVQPNQGRVTLTGELGTTQACNCSIRGKVTQVPSQHM 300
 QY 298 WQONLLQLTYPNGASYDPTDEVPAPLCTQDFSGMLYGVLTQDNVNVSTGEAKNAGIYIS 357
 DB 301 WN---LEITNLNGTQDFDTPDPAFLGVDFDFAFEGVGLSQNRNCS-NPANRAHDAVVA 356
 QY 358 TTSKGFTPKIGISGLHSITEHVHPNQOSRFTPVGV-AVDENTPFQOQVLPVHAGSLALNT 416
 DB 357 TYSKDTKPLGLVQGTWNTNDVENQPTKFTFIFGLNEVANGHRFQWTLPRYSGALTNNM 416
 QY 417 NLAPAVAPTPFGEQLLFFRSVPCVQGLQGDFAIDCLLPQSWNVHFTQEAAPSQADVAL 476
 DB 417 NLAPAVAPLFGERLLFFRSYVP-LKGGFGNPA-IDCSVPQEWQHFYQESAPSLGDVAL 474
 QY 477 IRYVNPDTGRTLFEAKLHRSAGITVSHTCAYPLVPPNGHFRFDSWVNFQFYSLAPMGTCN 536
 DB 475 VRYVNPDTGRTLFEAKLHRSAGITVSHTCAYPLVPPNGHFRFDSWVNFQFYSLAPMGTCN 534

QY 537 GRRRIQ 542
 DB 535 GRRRVQ 540
 RESULT 4
 AAB49705
 ID AAB49705 standard; protein; 548 AA.
 XX
 AC AAB49705;
 XX
 DT 04-APR-2001 (first entry)
 XX
 DE Small round structured virus protein SEQ ID 6.
 XX
 KW Small round structured virus; SRSV; food poisoning.
 XX
 OS Small round structured virus.
 XX WO200079280-A1.
 XX
 XX 28-DEC-2000.
 XX
 XX 22-JUN-2000; 2000WO-JP004095.
 XX
 XX 22-JUN-1999; 99JP-00175928.
 XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.
 XX
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 XX WPI; 2001-080848/09.
 DR N-PSDB; AAF29146.
 XX
 XX Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX
 XX Claim 1; Page 52-54; 84pp; Japanese.
 XX
 CC This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 CC
 XX Sequence 548 AA;
 SQ
 Query Match 67.4%; Score 1931; DB 4; Length 548;
 Best Local Similarity 65.7%; Pred. No. 3.9e-157; Mismatches 102; Indels 18; Gaps 5;
 Matches 364; Conservative 70;
 QY 1 MKWASNDAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQTNIIIDPWIRTNFVQAP 60
 DB 1 MKWASNDAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQNIIDPWIRTNFVQAP 60
 QY 61 NGEFTVSPRNSGCEILLNLELGPDLNPLYLAHLSRMVNGYAGGVQVLLAGNAFTAGKIL 120
 DB 61 GGEFTVSPRNSGCEILLNLELGPDLNPLYLAHLSRMVNGYAGGVQVLLAGNAFTAGKII 120
 QY 121 FAAPPPFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNPFVHNNQPPQMRLLVA 180
 DB 121 FAAPPPFPFVNLSPSQITMFPVHVIDVRLTLEPVLPMVDVRSSTLTFHFNQDEPKRLLVA 180
 QY 181 MLYTPLRNSGDDVFTVSCRVLTPTDPFETIYLVPPSVESKTKPFTLPILTITSELNLS 240
 DB 181 MLYTPLRNSGDDVFTVSCRVLTPTDPFETIYLVPPSVESKTKPFTLPILTITSELNLS 240
 QY 241 RFPPIEQLYTAPNETNVVQCNQGRCTLDGELQGTQTLSSAVCFLOG-----RTVADN 294

PS Claim 1; Page 62-64; 84pp; Japanese.

XX This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX Sequence 550 AA;

Query Match 65.4%; Score 1872; DB 4; Length 550;
 Best Local Similarity 64.9%; Pred. No. 4.7e-152;
 Matches 361; Conservative 74; Mismatches 101; Indels 20; Gaps 7;

QY 1 MKMASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQNIIDPWIRTFVQAP 60
 DB 1 MKMASNDAAPSNDGAANLVPEANDEVMALFVVGASIAAPVVGQNIIDPWIRENFVQAP 60

QY 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGGVEVQVLLAGNAFTAGKIL 120
 DB 61 QGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGGVEVQVLLAGNAFTAGKII 120

QY 121 FFAAIPNPLVDMSIAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNQPPMRLLVA 180
 DB 121 FFAAIPNPLVDMSIAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNQPPMRLLVA 180

QY 181 MLYTPLRSNGSGDDVFTVSCRVLTPTDPEFIYLVPSVESKTPFTLPILTISELTNS 240
 DB 181 MLYTPLRSNGSGDDVFTVSCRVLTPTDPEFIYLVPSVESKTPFTLPILTISELTNS 240

QY 241 RPPPIEQLYTAPNETNVVQCGRCCTLDGELQGTTLSSAVCFVQGRVADN---GN 297
 DB 241 RPPPIEQLYTAPNETNVVQCGRCCTLDGELQGTTLSSAVCFVQGRVADN---GN 297

QY 298 WQNLQLQTYP-----NGASYDPTDEVPAPLGTQDFPSGMLYGVLTQDNVNVSTGE-- 347
 DB 298 WQNLQLQTYP-----NGASYDPTDEVPAPLGTQDFPSGMLYGVLTQDNVNVSTGE-- 347

QY 348 -AKNAKGIYSTSKFTPKIGSLHSITHEVHNQSRFTPVGVAVDENTPFOOWVLP 406
 DB 348 -AKNAKGIYSTSKFTPKIGSLHSITHEVHNQSRFTPVGVAVDENTPFOOWVLP 406

QY 407 HYAGSLALNTLAPAVTPFGEQQLFFRSRVPVCGVQGLQGDQDAFIDCLLPQEWNVHFOE 466
 DB 407 HYAGSLALNTLAPAVTPFGEQQLFFRSRVPVCGVQGLQGDQDAFIDCLLPQEWNVHFOE 466

QY 467 AAPSQADVALIRYVNPDTGRTLFEAKLHRSFGITVSHGTGAYPLVPPNNGHFRFDSWVNOF 526
 DB 467 AAPSQADVALIRYVNPDTGRTLFEAKLHRSFGITVSHGTGAYPLVPPNNGHFRFDSWVNOF 526

QY 527 YSLAPMGTCNGRRRQ 542
 DB 527 YSLAPMGTCNGRRRQ 542

QY 535 YTLAPMGSCQGRRAQ 550
 DB 535 YTLAPMGSCQGRRAQ 550

RESULT 7
 AAB49704
 ID AAB49704 standard; protein; 539 AA.
 XX
 XX AAB49704;
 AC
 DT 04-APR-2001 (first entry)
 XX
 DE Small round structured virus protein SEQ ID 5.
 XX
 KW Small round structured virus; SRSV; food poisoning.
 XX
 OS Small round structured virus.
 XX
 PN WO200079280-A1.
 XX

PD 28-DEC-2000.

XX 22-JUN-2000; 2000WO-JP004095.

XX 22-JUN-1999; 95JP-00175928.

XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

PA (DENK-) DENKA SEIKEN KK.

PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

XX MPI; 2001-080848/09.

DR N-PSDB; AAF29145.

XX Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.

XX Claim 1; Page 50-52; 84pp; Japanese.

CC This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX Sequence 539 AA;

Query Match 63.6%; Score 1821.5; DB 4; Length 539;
 Best Local Similarity 63.7%; Pred. No. 1e-147;
 Matches 347; Conservative 74; Mismatches 111; Indels 13; Gaps 6;

QY 1 MKMASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQNIIDPWIRTFVQAP 60
 DB 1 MKMASNDAAPSNDGAANLVPEANDEVMALFVVGASIAAPVVGQNIIDPWIRENFVQAP 60

QY 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGGVEVQVLLAGNAFTAGKIL 120
 DB 61 QGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGGVEVQVLLAGNAFTAGKII 120

QY 121 FFAAIPNPLVDMSIAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNQPPMRLLVA 180
 DB 121 FFAAIPNPLVDMSIAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNQPPMRLLVA 180

QY 181 MLYTPLRSNGSGDDVFTVSCRVLTPTDPEFIYLVPSVESKTPFTLPILTISELTNS 240
 DB 181 MLYTPLRSNGSGDDVFTVSCRVLTPTDPEFIYLVPSVESKTPFTLPILTISELTNS 240

QY 241 RPPPIEQLYTAPNETNVVQCGRCCTLDGELQGTTLSSAVCFVQGRVADN---GN 298
 DB 241 RPPPIEQLYTAPNETNVVQCGRCCTLDGELQGTTLSSAVCFVQGRVADN---GN 298

QY 301 NLLQTYPNNGASYDPTDEVPAPLGTQDFPSGMLYGVLTQDNVNVSTGEAKNAKGIYSTTS 360
 DB 301 NLLQTYPNNGASYDPTDEVPAPLGTQDFPSGMLYGVLTQDNVNVSTGEAKNAKGIYSTTS 360

QY 355 YTMNLASQNWSDYDTEIPAPLGTDPVGIQGLMTQTTRDGDGSTRHAKT---VSTGS 355
 DB 355 YTMNLASQNWSDYDTEIPAPLGTDPVGIQGLMTQTTRDGDGSTRHAKT---VSTGS 355

QY 361 CKFTPKIGSLHSITHEVHNQSRFTPVGVAD----ENTPQOQVLPVHAGSLALN 415
 DB 361 CKFTPKIGSLHSITHEVHNQSRFTPVGVAD----ENTPQOQVLPVHAGSLALN 415

QY 416 TNLAPAVAPTFFGEQQLFFRSRVPVCGVQGLQGDQDAFIDCLLPQEWNVHFOEAAAPSQADVA 475
 DB 416 TNLAPAVAPTFFGEQQLFFRSRVPVCGVQGLQGDQDAFIDCLLPQEWNVHFOEAAAPSQADVA 475

QY 476 LIRYVNPDTGRTLFEAKLHRSFGITVSHGTGAYPLVPPNNGHFRFDSWVNOFYSLAPMGTC 535
 DB 476 LIRYVNPDTGRTLFEAKLHRSFGITVSHGTGAYPLVPPNNGHFRFDSWVNOFYSLAPMGTC 535

QY 536 NGRRR 540
 DB 536 NGRRR 540

QY 533 AGRRR 537
 DB 533 AGRRR 537

QY 56 FVQAPNGEFTVSPRNSPGBIILNLELGPDLNPLYLHLRSMYNGYAGGVEVQVLLAGNAPT 115
DB 60 FVQAPNGEFTISPNNTPGGVLFDSLGLPHLNPFLHLRSMYNGVGNMRVRLMAGNAPT 119
QY 116 AGKILFAIPNPLNPLDMSPAQITMLPHLIVDVRTLEPMTPLDVRNVFVHFNQPPQR 175
DB 120 AGKIIVSCIPPGFGSHNLITIAQATLFPFHVIADVRTLDPIEVLDPVNRNVLFFHNNRNOOT 179
QY 176 MELVAMLYTLPLRSNGSGDDVFTVSCRVLTRPTDPDEFIYLVPPSVESKTKPFTLPILITIS 235
DB 180 MELVCMLYTLPLRTGGTGDSFVAGRVMTCPSPDNFLFLVPTVEQKTRPFTLPNPLS 239
QY 236 ELTNSRFPPIPIEQLYTAPNETNVQCGRCRTLDGELQGTQTLSSAVCFLOGRTVADNG 295
DB 240 SLNSRAPLPISSMGISPDNVQSFQNGRCRTLDGRLVGTTPVSLSHVAKIRG---TSNG 296
QY 296 DNWDQNLQLYTPNGASVDPDTEVPAPLGTQDFSGMLYGVLTQDNNVNVSTGEAKNAGIY 355
DB 297 -----TVINLTDELDTGTPHFF-EGPAPIGFPLDGGCDW-----HINMTQFGHSSQTYD 344
QY 356 ISTTSKFTPKIGSIHLHSITEHVHPNQSRFTFVGVAVDENTPFQOQVLP 407
DB 345 VDTTPTDTPVPHLGSIQANGISGNVGVLSWSPESH-----PSGQVD-----LWKIPN 394
QY 408 YAGSLALNTNLAPAVAPTPPGEQQLFFRSRVPVCGLOQODAF-IDCLLPQEWVNHFYOE 466
DB 395 YGSSITEATHLAPSVVPPGFGVLFVFFMSKIP-----GPGAYSLPCLLPQYISHLASE 448
QY 467 AAPSOADVALIRYVNPDTGRTILFEAKLHRSGETITVSHTCAY--PLVVPNGHFRPDSWN 524
DB 449 QAPTVEAALLHYVDPDTGRTILGEFKAYPDGFLTCVPNGASSGPOQLPINGVVFVSVWS 508
QY 525 QFYLAPMGITGNRR 539
DB 509 RPYQLKPVGTASSAR 523

RESULT 10

AARS0972

ID AARS0972 standard; protein; 530 AA.

XX AC AARS0972;

XX DT 16-OCT-2003 (revised)

XX DT 25-MAR-2003 (revised)

XX DT 05-OCT-1994 (first entry)

XX DE Norwalk virus strain 8FIIa protein (encoded by ORF2).

XX KW Norwalk virus; pathogen; acute gastroenteritis; food poisoning;

XX KW seafood contamination; diagnostic assay; calicivirus; small round virus.

XX OS Norwalk virus; (strain 8FIIa).

XX PN WO9405700-A2.

XX PD 17-MAR-1994.

XX PF 07-SEP-1993; 93WO-US008447.

XX PR 07-SEP-1992; 92US-00941365.

XX PA (BAYU) BAYLOR COLLEGE MEDICINE.

XX PI Matson DO, Estes MK, Jiang X, Graham DY;

XX DR WPI; 1994-101125/12.

XX DR N-PSDB; AAQ56826.

XX PT DNA from Norwalk and related viruses - used for preparing prods. for use

XX PT in diagnostic assays, detection and vaccines for Norwalk and related

XX PT viruses.

XX

PS Claim 14; Page 68-70; 156pp; English.

XX CC The Norwalk virus was isolated from stool samples from adult volunteers

CC infected with safety tested Norwalk virus strain 8FIIa. The coding

CC sequence is useful for the design of probes for use in diagnostic assays

CC for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct

CC PN field.) (Updated on 16-OCT-2003 to standardise OS field)

XX SQ Sequence 530 AA;

Query Match 41.2%; Score 1179; DB 2; Length 530;

Best Local Similarity 46.5%; Pred. No. 2.1e-92;

Matches 256; Conservative 84; Mismatches 171; Indels 40; Gaps 12;

QY 1 MKMASDAPNSNDGAA---SLVPEGINETMPL--BPVAGASIAAIVAGOTNIIDPWIRN 55

DB 1 MMASKDATSSVDGASGAGQLVPE-VNASDPLAMPDVAGSSTAVATAGQVNIIDPWIRN 59

QY 56 FVQAPNGEFTVSPRNSPGBIILNLELGPDLNPLYLHLRSMYNGYAGGVEVQVLLAGNAPT 115

DB 60 FVQAPNGEFTISPNNTPGGVLFDSLGLPHLNPFLHLRSMYNGVGNMRVRLMAGNAPT 119

QY 116 AGKILFAIPNPLNPLDMSPAQITMLPHLIVDVRTLEPMTPLDVRNVFVHFNQPPQR 175

DB 120 AGKIIVSCIPPGFGSHNLITIAQATLFPFHVIADVRTLDPIEVLDPVNRNVLFFHNNRNOOT 179

QY 176 MELVAMLYTLPLRSNGSGDDVFTVSCRVLTRPTDPDEFIYLVPPSVESKTKPFTLPILITIS 235

DB 180 MELVCMLYTLPLRTGGTGDSFVAGRVMTCPSPDNFLFLVPTVEQKTRPFTLPNPLS 239

QY 236 ELTNSRFPPIPIEQLYTAPNETNVQCGRCRTLDGELQGTQTLSSAVCFLOGRTVADNG 295

DB 240 SLNSRAPLPISSMGISPDNVQSFQNGRCRTLDGRLVGTTPVSLSHVAKIRG---TSNG 296

QY 296 DNWDQNLQLYTPNGASVDPDTEVPAPLGTQDFSGMLYGVLTQDNNVNVSTGEAKNAGIY 355

DB 297 -----TVINLTDELDTGTPHFF-EGPAPIGFPLDGGCDW-----HINMTQFGHSSQTYD 344

QY 356 ISTTSKFTPKIGSIHLHSITEHVHPNQSRFTFVGVAVDENTPFQOQVLP 411

DB 345 VDTTPTDTPVPHLGSIQANGISGNVGVLSWSPESH-----LWKIPN 398

QY 412 LALNTNLAPAVAPTPPGEQQLFFRSRVPVCGLOQODAF-IDCLLPQEWVNHFYOE 470

DB 399 ITEATHLAPSVVPPGFGVLFVFFMSKIP-----GPGAYNLPCLLPQYISHLASEQAPT 452

QY 471 QADVALIRYVNPDTGRTILFEAKLHRSGETITVSHTCAY--PLVVPNGHFRPDSWNQFVS 528

DB 453 VGEAALLHYVDPDTGRTILGEFKAYPDGFLTCVPNGASSGPOQLPINGVVFVSVWSRFFQY 512

QY 529 LAPMGITGNRR 539

DB 513 LKPVGTASSAR 523

RESULT 11

ADC72176

ID ADC72176 standard; protein; 530 AA.

XX AC ADC72176;

XX DT 18-DEC-2003 (first entry)

XX DE Norwalk virus protein 2 amino acid sequence.

XX KW immune response; non-Norwalk virus agent; immunogen; Norwalk virus;

XX KW viral pathogen; acute gastroenteritis; virucidal; antiviral vaccine.

XX OS Norwalk virus.

XX PN US6572862-B1.

XX

```

PD 03-JUN-2003.
XX
PF 07-JUN-1995; 95US-00486049.
XX
XX 08-NOV-1989; 89US-00433492.
XX
PR 27-APR-1990; 90US-00515993.
PR
PR 27-AUG-1990; 90US-00573509.
PR
PR 06-MAY-1991; 91US-00696454.
XX
XX (BAYU ) BAYLOR COLLEGE MEDICINE.
PA
XX
XX Estes MK, Jiang X, Graham DY;
PI
XX
XX WPI; 2003-776005/73.
DR
DR N-PSDB; ADC72174.
XX
XX Inducing an immune response against non-Norwalk virus agents, comprises
PT administering an immunogen recombinantly expressed from a cDNA from
PT Norwalk virus.
XX
XX
XX Example 4; SEQ ID NO 3; 45pp; English.
XX
XX This invention relates to a novel method of inducing an immune response
XX in an individual against Norwalk virus and non-Norwalk virus agents, by
XX orally or parenterally administering an immunogen recombinantly expressed
XX or synthesised from a cDNA of Norwalk virus given in the specification.
XX Norwalk virus is one of the most important viral pathogens, causing acute
XX gastroenteritis. The invention may be used for the development of
XX compounds with virucidal activity or an antiviral vaccine. The present
XX sequence is the amino acid sequence of a protein encoded by the Norwalk
XX virus genome of the invention.
XX
XX
XX Sequence 530 AA;
XX
XX
XX Query Match 41.2%; Score 1179; DB 7; Length 530;
XX Best Local Similarity 46.5%; Pred. No. 2.1e-92;
XX Matches 256; Conservative 84; Mismatches 171; Indels 40; Gaps 12;
XX
XX 1 MKMASNDAAPSNDGAA---SLVPGINETWPL--EPVAGASIAAPVAGOTNIIDPWRTN 55
XX 1 MMMAASKDATSSVDGASGAGQVPE-VNASDPLAMDVPAGSTAVATAGQNPIDPWRTN 59
XX
XX 56 FVOAPNGEFTVSPRNSPGEILLNLELGPDLNPLLAHLSRMVYAGGVQVLLAGNAFT 115
XX 60 FVOAPQGEFTISPNTPGDVLFDSLGLPHLNPFLHLSQMTNGVGNVRIMLAGNAFT 119
XX
XX 116 AGKILFAAIPNPLVDMTSPAQITMLPHLIVDVTLEPIMTLPDVRNVFHHNQPPR 175
XX 120 AGKIIVSCIPPGFGSHNLTAQATLFPFHVIADVRLDPIEVPLEDVRNVLFHNNDRNQOT 179
XX
XX 176 MRLVAMLYTPLRNSGDDVFTVSCRVLTRPTDPPEFIYLVPPSVESKTKPFTLILTIS 235
XX 180 MRLVCMLYTPLTGGTGDSFVAGRWTCSPDPENFLFVPTVEQTRFPTLPNPLS 239
XX
XX 236 ELTNSRFPIPIQLYATARNVTVQNGRCRTLDGELQGTTLQSSAVCFLOQRTVADNG 295
XX 240 SLNSRAPLPISSMGISPDNVQSQFQNGRCRTLDGLVGTTPVSLSHVAKIRG---TSNG 296
XX
XX 296 DNWDQNLQLTYPNGASYDPTDEVAPLGTQDFSGMLXGVLTDQDNVNVSTGEAKNAGIY 355
XX 297 -----TVINLTDLGTPPEHP-EGPAPIGPDLGGCDM-----HINTQFCHSSQTYD 344
XX
XX 356 ISTTSCKFTPKIGSLGHSITEHVHVNQOSRFT----PVGVAVENTFPQOVLPHYAGS 411
XX 345 VDTTPTTFVPHLGSIOANGISGVYGVLSWISPSHPSGSQVD-----LWKIPNYGSS 398
XX
XX 412 LALNTLAPAVAPTFPGEOLLFPRSRVPCVQLOGODAF-IDCLLPQEWNVHFYEAAPS 470
XX 399 ITEATHLAPSVYPGGEVLVFFMFSKMP-----GPGAYNLCPQLPQETISHLASEQAPT 452
XX
XX 471 QADVLRVYNPDGTGRTLPEAKLHRGGFTIVTSHTGAY--PLVVPNGHFRFPDSWNQFYS 528
XX 453 VGEAALLHYVDPTGRTNLGFEKAYPDGFLTCVNGASSGPQQLPFGVVFVSVSRFYQ 512

```

```

QY 529 LAPMGTGNGRR 539
DB 513 LKPVGTASSAR 523

RESULT 12
AAB49700
ID AAB49700 standard; protein; 545 AA.
XX
XX AAB49700;
XX AC
XX 04-APR-2001 (first entry)
XX
XX Small round structured virus protein SEQ ID 1.
XX
XX Small round structured virus; SRSV; food poisoning.
XX
XX Small round structured virus.
XX
XX WO200079280-A1.
XX
XX 28-DEC-2000.
XX
XX 22-JUN-2000; 2000WO-JP004095.
XX
XX 22-JUN-1999; 99JP-00175928.
XX
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI; 2001-080848/09.
XX N-PSDB; AAF29141.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX strains for investigation of food poisoning outbreaks, contains
XX antibodies.
XX
XX Claim 1; Page 40-42; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710, -
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 545 AA;
XX
XX Query Match 41.0%; Score 1172.5; DB 4; Length 545;
XX Best Local Similarity 46.4%; Pred. No. 8e-92;
XX Matches 257; Conservative 85; Mismatches 185; Indels 27; Gaps 15;
XX
XX 1 MKMASNDAAPSN---DGAASLVPE-GINETWPLEFPVAGASIAAPVAGOTNIIDPWRTN 55
XX 1 MMMAASKD-APTNWDGTSGAGQLVPEANTAEPISEMFVAGATAATAATAGQVNMIDPWIMN 59
XX
XX 56 FVOAPNGEFTVSPRNSPGEILLNLELGPDLNPLLAHLSRMVYAGGVQVLLAGNAFT 115
XX 60 FVOAPQGEFTISPNTPGDVLFDSLGLPHLNPFLHLSQMTNGVGNVRIMLAGNAFT 119
XX
XX 116 AGKILFAAIPNPLVDMTSPAQITMLPHLIVDVTLEPIMTLPDVRNVFHHNQPPR 175
XX 120 AGKIIVSCIPPGFGSHNLTAQATLFPFHVIADVRLDPIEVPLEDVRNVLFHNNDRNQOT 178
XX
XX 176 MRLVAMLYTPLRNSG--SGDDVFTVSCRVLTRPTDPPEFIYLVPPSVESKTKPFTLILT 233
XX 179 MRLVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVFPNVQKTFPFSVFNLP 238
XX 234 ISELTNSRFPPIPIQLYATARNVTVQNGRCRTLDGELQGTTLQSSAVCFLOQRTVAD 293

```

```

Db 239 LMTLSNRVPSLIKSMWRDQGMVQFQNGRVTLTGQDGGTTPTASOLCKIRGSVFHA 298
Qy 294 NGDNWDQNLQLTYPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAG 353
Db 299 NGGN-GYNLT- --DGSFYH-AFESPAIGPFDLGECDHMEASFTTQFNTGDV--IKQ 351
Qy 354 IYISTSGKFTPKIGSI-----GLHSITEHVHPNQQRFTPVGVAVD-ENTPFOQWLVPHA 409
Db 352 INVQESA-FAPHLGTIQADGLSDVS--VNTNIAKLGWVSPVSDCHRGDVPWIPRYG 408
Qy 410 GSLALNTNLAPAVPTFPCEQLLFRSRVPCVQGLQGQDAFIDCLLPQSMWVHFQEARP 469
Db 409 SILTEAAQLAPPIYPPGFGAIVFFMSDFPIAHGNTGLS--VFCTIPQEFVTHFVNEQAP 466
Qy 470 SQADVALIRYVNDTGRITFEAKLHRSGETIV--SHTGAYPLVVPNGHFRDWSWNOFY 527
Db 467 TRGEAALLHYLPDTHNLGEFKLYPEGEFTWCVPNSSGTGPQLPLINGVVFVFSWSRY 526
Qy 528 SLAPMGTGNRRRI 541
Db 527 QLKPVGTAGPACRL 540

```

RESULT 13

```

ID AAB49703
XX AAB49703 standard; protein; 544 AA.
AC AAB49703;
XX
DT 04-APR-2001 (first entry)
DE Small round structured virus protein SEQ ID 4.
KW Small round structured virus; SRSV; food poisoning.
XX
OS Small round structured virus.
XX
PN WO200079280-A1.
XX
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000WO-JP004095.
XX
PR 22-JUN-1999; 99JP-00175928.
XX
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
XX
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
DR WPI: 2001-080848/09.
DR N-PSDB; AAF29144.
XX
PT Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
PS Claim 1; Page 47-49; 84pp; Japanese.
XX
CC This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX
SQ Sequence 544 AA;

```

Query Match 40.7%; Score 1165.5; DB 4; Length 544;

Best Local Similarity 45.3%; Pred. No. 3.2e-91;

Matches 256; Conservative 89; Mismatches 171; Indels 49; Gaps 16;

```

Qy 1 MQGAANDRAAPSND---GAASLVPEGIN--ETWPLEBPVAGASTAAPVAGGTNIIDPWIRN 55
Db 1 MMASKDATPSADGATGAGQLVPE-VNTADPIPIDVAGSGSTALATAGQNLIDPMIINN 59
Qy 56 FVQAPNGEFTVSPRNSPGSEILLNLNLELGDPLNLYLAHLSRMVNGYAGGVEQVLLAGNAPT 115
Db 60 FVQAPQGBFTISPNNTPGDVLPDLQGLPHLNFPLSHLSQMYNGWGNMKRVVVLGNAPT 119
Qy 116 AGKILFAAIPNPNFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVVYHFNNOQPR 175
Db 120 AGKVIICCVPCFQFQSRITLSIAQATLFPHVIAADVRLTDPVEVEPLEDVRNVLYH-NNDTQT 178
Qy 176 MRLVAMLYTPLRNSGS--GDDVFTVSCRVLTRPTPDFEYIYLVPPSVSCKTPFTLPILT 233
Db 179 NELLCLWLYTPLRTGASGTDTSFVAVGRVLTCPGPDFNPLFLVPPTEQKTRPFTVPNIP 238
Qy 234 ISELNTRPPPIEOLYTPAPNETNVVQNGRECTLDGELQGTQTLSSAVCFLOGRTVAD 293
Db 239 LKYLNSRIPNPIEGMSLSPDQTNVQFQNGRCTIDGQPLGTTPVSVSOLCKFRGRTS- 297
Qy 294 NGDNWDQNLQLTYPNGASYDPTDEVPAPLGTQDFSGMLYGV-----LTQDNVNVST 345
Db 298 -----GQVNLNLTLDGSPF-WAPAAPAPAGFPDGLSCDWHIEMSKIPNSSTQNNPIVTN 351
Qy 346 GEAKNAKGIYISTSGKFTPKIGSIGL----HSITEHVHPNQQRFTPVGVAVDENTPFO 401
Db 352 S-----VKPNSQQFVPHLSITLDBENVSSGCGYIGTIQWT--SPSPDSGGANTNF- 399
Qy 402 QWVLPHYAGSLALNTNLAPAVAPTFPGEQLLFRSRVPCVQGLQGQDAFIDCLLPQSMWV 461
Db 400 -WKIPDYGSSLAESAASQLAPAVYPGNEVIVYFMAVIP-GPNQSGSPNLVPCLLPQEYIT 457
Qy 462 HFYQEAAPSADVALIRYVNDTGRITLFEAKLHRSGETIV-----SHTGAYPLVVPNGH 517
Db 458 HFISEQAFIQGEAALLHYVDPTNRLNLEGEFKLYPCGYLTCVPNSSSTG--PQQLPLDG 515
Qy 518 RFDWSWNOFYSLAPMGT-GNGRRRI 541
Db 516 VFASWVSRYQLKPVGTAGPARGRL 540

```

RESULT 14

```

ID AAB49702
XX AAB49702 standard; protein; 546 AA.
AC AAB49702;
XX
DT 04-APR-2001 (first entry)
DE Small round structured virus protein SEQ ID 3.
KW Small round structured virus; SRSV; food poisoning.
XX
OS Small round structured virus.
XX
PN WO200079280-A1.
XX
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000WO-JP004095.
XX
PR 22-JUN-1999; 99JP-00175928.
XX
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
XX
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
DR WPI: 2001-080848/09.
DR N-PSDB; AAF29143.
XX
PT Kit for the detection and typing of small round-structured virus (SRSV)

```

Wed Jun 2 09:13:34 2004

us-09-926-799-9.rag

strains for investigation of food poisoning outbreaks, contains antibodies.

Claim 1; Page 45-47; 84pp; Japanese.

This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

Sequence 546 AA;

Query Match 40.4%; Score 1155.5; DB 4; Length 546;
Best Local Similarity 45.3%; Pred. No. 2.3e-90;
Matches 258; Conservative 88; Mismatches 171; Indels 53; Gaps 18;

QY 1 MKMASNDAAPSNDGAA--SLVPEGIN--ETWPLEPVAGASIAAPVAGQTNIIDPWRTN 55
DB 1 MMWASKDAPQSDAGSAGAGQLVPE-VNTADPLPMEVPVAGFTTAVATAGOVNMIDPWVNN 59
QY 56 FVOAPNGEFTVSPRNSPGEILLNLELGPDLNPLVLAHLSRMVNGYAGVEVQVLLAGNAFT 115
DB 60 FVQSPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGNMRVILLAGNAPS 119
QY 116 AGKILFAAIPNPLVDMSIPAQITMLPHLIVDVRTLEPMTPLDVRNVYFHNQOPQR 175
DB 120 AGKIIVCCVPPGFTSSLLTQAATLPHVIAVDTLEPIEMPLEDVRNVLYHTNDN-QPT 178
QY 176 MRLVAMLYTLPLRS-NGSGD-DVFTVSCRVLTRTPDFEFYLYVPPSVESKTKPFTLPILT 233
DB 179 MRLVCMLYTLPLRTGGSGNSDSFVWAGRVLTAPSSDFSFLFLVPEPTIEQKRAFTVFNIP 238
QY 234 ISELTNSRFPPIEQLYTAPEVNTVQNGRCTLDELQGTQALLSSAVCFQGRTRVAD 293
DB 239 LQTLNSRFPPIEQLYTAPEVNTVQNGRCTLDELQGTQALLSSAVCFQGRTRVAD 295
QY 294 NGDNWDQNLQLTYPNGASVDPDEVPAPLQGTQFSGMLYGVLTQDNVNVSTGBAKNAK 352
DB 296 ---NQARTNLNTEVDGKFFWAPDS-PAPVGFDPFGKCDWHMRISKTENNNTSSGDP--MR 349
QY 353 GIYISTSGKFTPKIGSI---GLHS-----NQOSRETPVGVAVDENTPFQ 401
DB 350 SVSVQTNVQGVFPHLGSIQFDEVFN--HPTGDYIGTIEWISQPS---TPPG-----TDIN 399
QY 402 QWVLPHYAGSLALNTNLAPVAPTGGQLLFFRSRVPVCGVQLOGQDA--FIDCLLPQEW 459
DB 400 LWEIPDYGSSLSQAANLAPVFPFGFGEALYFVSAPB---GPNRSAPNDVPCLLPQY 456
QY 460 VNHFYQEAAPSQADVALTRYNPDTGRTLFEAKLHRSQFIV--SHTGAYPLVVPNGHF 517
DB 457 ITHFVSEQAPTMDGAALLHYVDPTNRLGEBFKLYPGGYLTCVPGVAGPQQLPLNGVF 516
QY 518 RFDVSNVQFSLAPMGTGN-----GRRRI 541
DB 517 LFFVSWVSRYQLKPGVGTABTARSRLQVRRI 546

RESULT 15

AAR57091 standard; protein; 530 AA.

XX AAR57091;

XX 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 05-OCT-1994 (first entry)

XX Small round virus SRSV/KY/89 capsid protein.

XX pathogen; acute gastroenteritis; food poisoning; seafood contamination;

diagnostic assay; human calicivirus; small round virus; SRSV; KY89;
Norwalk virus; capsid protein.

Small round structured virus.

WO9405700-A2.

17-MAR-1994.

07-SEP-1993; 93WO-US008447.

07-SEP-1992; 92US-00941365.

(BAYU) BAYLOR COLLEGE MEDICINE.

Matson DO, Estes MK, Jiang X, Graham DY;

WPI; 1994-101125/12.

N-PSDB; AAQ56832.

DNA from Norwalk and related viruses - used for preparing prods. for use in diagnostic assays, detection and vaccines for Norwalk and related viruses.

Example 7; Fig 13a; 156pp; English.

The known sequence for Norwalk virus was used to obtain the sequence of other Norwalk-related viruses such as SRSV/KY/89, an agent from a stool from an outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide cDNA sequence includes part of the polymerase region and the capsid region of the genome; the deduced amino acid sequences are AAR57092 and AAR57091, respectively. Expression of fragments and derivs. of Norwalk-related viruses permits development of diagnostic assays to detect antibodies, antigens, viral genetic material or antivirals. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS field.)

Sequence 530 AA;

Query Match 39.7%; Score 1137; DB 2; Length 530;

Best Local Similarity 45.8%; Pred. No. 8.7e-89;

Matches 254; Conservative 85; Mismatches 168; Indels 48; Gaps 16;

QY 1 MKMASNDAAPSNDGAA--LVPEGINETMPL--EPVAGASIAAPVAGQTNIIDPWRTN 55
DB 1 MMWASKDATSVDGASASVQLVPE-VNASDPLAMDVPAGSSTAVATAGOVNPDWVNN 59
QY 56 FVOAPNGEFTVSPRNSPGEILLNLELGPDLNPLVLAHLSRMVNGYAGVEVQVLLAGNAFT 115
DB 60 FVOAPQGEFTISPNTPGDVLFDLSGLPHLNPFLSHLSQMYNGWGNMRVIMLAGNAFT 119
QY 116 AGKILFAAIPNPLVDMSIPAQITMLPHLIVDVRTLEPMTPLDVRNVYFHNQOPQR 175
DB 120 AGKIIVCCVPPGFGSQQLTQAATLPHVIAVDTLEPIEMPLEDVRNVLYHTNDNQQT 179
QY 176 MRLVAMLYTLPLRSNGSGDDVFTVSCRVLTRTPDFEFYLYVPPSVESKTKPFTLPILTIS 235
DB 180 MRLVCMLYTLPLSTGGTGDSFVWAGRVMTCPSPDENFLVLPVTEQKTRPFTLPENLPLS 239
QY 236 ELTNSRFPPIEQLYTAPEVNTVQNGRCTLDELQGTQALLSSAVCFQGRTRVADNG 295
DB 240 SLSNSRAPLPSGMSISPDNVQSVQNGRCTLDELQGTQALLSSAVCFQGRTRVADNG 296
QY 296 DNDWDQNLQLTYPNGASVDPDEVPAPLQGTQFSGMLYGVLTQDNVNVSTGBAKNAK 355
DB 297 ---TVINLTDELDPFPHF-EGPAPIGFDPDLGGCDW-----HINMTQFGHSSQTYD 344
QY 356 ISTTSGKFTPKIGSI---GLHS-----ITEHVHNPQOSRFTPVGVAVDENTPFQWVLPH 407
DB 345 VDTTPDTSVPHLGSIQANGIGSGNYIGVLSWSPSPSH---PSGSQVD-----LWKIN 394
QY 408 YAGSLALNTNLAPV-APTFPGEQLLFFRSRVPVCGVQLOGQDAFIDCLLPQEWVNHFYOE 466

Db 395 YGSSITEATHLAPSVSPGF-GEVLVFFMSKIPG-----PGGDS-LPCLLPQGYISHLASE 448
Qy 467 AAPSQADVALIRYVNPDTGRTLF EAKLHRS GFITV SHTGAY--PLVVPENGHFRFDSWYN 524
Db 449 QAPTVEGEGFLLHYVDPTDNRNLGEFKAYPDGFLTCVPNGCASSGPQQLPINGVVFVSWYS 508
Qy 525 QFYSLAPMG TGNGRR 539
Db 509 RFYQLKPVGTASTAR 523

Search completed: June 1, 2004, 13:45:59
Job time : 47.5611 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 12.9134 Seconds
(without alignments)
2166.837 Million cell updates/sec

Title: US-09-926-799-9
Perfect score: 2863
Sequence: 1 MMASNDAPNDGASLVP.....VNQFYSLAPMTGNGRRRIQ 542

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA*
1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS-COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1179	41.2	530	4	US-08-486-049-3
2	278	9.7	623	4	US-09-590-020-2
3	278	9.7	623	4	US-09-590-020-4
4	275	9.6	622	4	US-09-590-020-6
5	272.5	9.5	626	4	US-09-590-020-7
6	265.5	9.3	668	4	US-09-617-594A-4
7	265.5	9.3	669	4	US-09-617-594A-2
8	115	4.0	2206	1	US-07-852-260-2
9	115	4.0	2206	2	US-08-461-503-2
10	115	4.0	2206	3	US-08-465-250-2
11	108.5	3.8	1091	6	5516630-2
12	108.5	3.8	2227	3	US-08-475-886-2
13	108.5	3.8	2227	3	US-08-475-886-4
14	108.5	3.8	2227	3	US-08-475-886-6
15	108.5	3.8	2227	3	US-08-397-232-2
16	108.5	3.8	2227	3	US-08-397-232-4
17	108.5	3.8	2227	3	US-09-171-387-2
18	108.5	3.8	2227	4	US-09-653-499-2
19	108.5	3.8	2227	4	US-09-653-499-4
20	108.5	3.8	2227	4	US-09-653-499-6
21	108.5	3.8	2227	4	US-10-104-966-12
22	108.5	3.8	2227	4	US-10-135-988-2
23	108.5	3.8	2227	4	US-10-135-988-4
24	108.5	3.8	2227	4	US-10-135-988-6
25	108	3.8	2318	3	US-09-091-219-24
26	108	3.8	2318	4	US-09-660-541-24
27	107.5	3.8	2736	4	US-09-252-991A-30227

28	106.5	3.7	740	4	US-10-164-595-6	Sequence 6, Appl1
29	106.5	3.7	755	4	US-10-164-595-10	Sequence 10, Appl1
30	106.5	3.7	781	4	US-10-164-595-4	Sequence 4, Appl1
31	104.5	3.7	762	4	US-09-107-532A-4275	Sequence 4275, Ap
32	103.5	3.6	839	1	US-08-087-016-2	Sequence 2, Appl1
33	101.5	3.5	1257	1	US-08-340-428B-49	Sequence 49, Appl1
34	100.5	3.5	496	4	US-10-164-595-8	Sequence 8, Appl1
35	100.5	3.5	496	4	US-10-164-595-60	Sequence 60, Appl1
36	99.5	3.5	1056	4	US-09-079-030-217	Sequence 217, App
37	99.5	3.5	4536	4	US-09-180-422B-27	Sequence 27, Appl
38	99.5	3.5	4536	4	US-09-079-030-1	Sequence 1, Appl1
39	99.5	3.5	4563	4	US-09-108-006C-1	Sequence 1, Appl1
40	98.5	3.4	1638	4	US-09-071-035-258	Sequence 258, App
41	98.5	3.4	1638	4	US-09-071-035-262	Sequence 262, App
42	98.5	3.4	1638	4	US-09-071-035-266	Sequence 266, App
43	96.5	3.4	579	6	5223424-13	Patent No. 5223424
44	95.5	3.3	834	4	US-09-252-991A-31321	Sequence 31321, A
45	95	3.3	426	4	US-09-107-532A-7172	Sequence 7172, Ap

ALIGNMENTS

RESULT 1
US-08-486-049-3
; Sequence 3, Application US/08486049
; Patent No. 6572862
; GENERAL INFORMATION:
; APPLICANT: Bates, Mary K
; APPLICANT: Jiang, Xi
; APPLICANT: Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Ave., N.W.
; CITY: Washington, D.C.
; STATE:
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,049
; FILING DATE: June 7, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311.023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-486-049-3

Query Match 41.2%; Score 1179; DB 4; Length 530;
Best Local Similarity 46.5%; Pred. NO. 5e-107;
Matches 256; Conservative 84; Mismatches 171; Indels 40; Gaps 12;
QY 1 MMASNDAPNDGAA---SLVPEGINETMPL--EPVAGASIAAPVAGQTHLIDPWITN 55
DB 1 MMASDKATSSVDGASGAGQLVPE-VNASDPLAMDVPVAGSTAVATAGQVNPIDPWITN 59

127 NPLVDMISPAQITML---PHLIVDVRTLEPIMTLPDVRNVFYHFNQPOQPMRLVAMLY 183
 194 G-----IEFVQSTMLQYPHVLFDFARQVEPVIFAIDRLSNLXHLMSD-TTSLVIMVY 247
 184 TPLRSNGSGDDVFTVSC--RVLTRPTDFEFYLVPP-----SVESK----- 223
 248 NDL-INPYANDTNSGGCIVTETKPGDFKPHLLKPPGSMITHGVSFSDLPKSSSLWIG 306
 224 -----TKPFTL-----PI-LTISELTNSRFPPIPIBQLY 250
 307 NRHWSIDITFIIRPFVFOANRHFDFNQTAGWSTPRFRPIITITVSESNMKLGIGVATDY 366
 251 TAPN-----ENW-----VOCQNGR--CTLDG-----ELQTTOLLSSAVCF 286
 367 IVGPIPDGPDPTIIEQLTTPAGIYSITASNGVITTAAGYDAAETIVNTTNFKSMYIC-- 424
 287 QGRTVADNDGNDQNLQIT--YPMGASYDPTDEV-----PAPLG---TODPSGM 331
 425 -GSLQRAWGDKKISNTAFITTAVRKNGSIEPSNTIDMTKLVVYQDAHVGEVQISDITLA 483
 332 LYGVLTQDNNVNVSTGEAKNAG-----IYISTTSCKFTPKIGSIGLHSHITEHVHPNOQS 385
 484 LGGY-----TGIGEEAIGSDRDKVVRISV---LPETGARGGN-----HPIFYK 523
 386 RFTPVGVAVDENTPFOQWVLPHVAGSLAINTLAP 420
 524 NSIKLGVIRSIDVFNSQL-HTSRQLSLNNYLLP 557

RESULT 3
 US-09-590-020-4
 ; Sequence 4, Application US/09590020
 ; Patent No. 6355246
 ; GENERAL INFORMATION:
 ; APPLICANT: Kruger, John M
 ; APPLICANT: Maes, Roger K
 ; APPLICANT: Vilnis, Aivars
 ; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
 ; TITLE OF INVENTION: CALCIVIRUS AND VACCINES THEREOF
 ; FILE REFERENCE: MSU 4.1-446
 ; CURRENT APPLICATION NUMBER: US/09/590,020
 ; CURRENT FILING DATE: 2000-06-08
 ; PRIOR APPLICATION NUMBER: 60/138,484
 ; PRIOR FILING DATE: 1999-06-10
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 623
 ; TYPE: PR1
 ; ORGANISM: Feline calicivirus
 ; US-09-590-020-4

Query Match 9.7%; Score 278; DB 4; Length 623;
 Best Local Similarity 25.0%; Pred. No. 2.6e-18;
 Matches 129; Conservative 65; Mismatches 177; Indels 144; Gaps 26;

12 NDGAASLVPEGINETMP-----LEPVAGASIAAPVAGQTNIIIDPW-IRTNFVQAPNGEFTV 66
 81 DDGSITTPEQG---TWVGGVIAEPGSAQMSAAADMATGKSDSEWEAFFSFHTSVNW---- 133
 67 SPRNSPEILLNLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKILFAAIPP 126
 134 STSETQKILFKQSLGPLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKILFAAIPP 193
 127 NPLVDMISPAQITML---PHLIVDVRTLEPIMTLPDVRNVFYHFNQPOQPMRLVAMLY 183
 194 G-----IEFVQSTMLQYPHVLFDFARQVEPVIFAIDRLSNLXHLMSD-TTSLVIMVY 247
 184 TPLRSNGSGDDVFTVSC--RVLTRPTDFEFYLVPP-----SVESK----- 223
 248 NDL-INPYANDTNSGGCIVTETKPGDFKPHLLKPPGSMITHGVSFSDLPKSSSLWIG 306
 224 -----TKPFTL-----PI-LTISELTNSRFPPIPIBQLY 250

56 FVOAPNGEFTVSPRNSPEILLNLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFT 115
 60 FVOAQSGEFTISPNTPGDVLFDLSGLPHLPFLHLSQMYNGWGNRMVRIMLAGNAFT 119
 116 AKKILFAAIPPNFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNQPOQPR 175
 120 AGKIIIVSCIPPGFGSHNTIAQATLFFHVIAADVRLTDFIEVPLEDVRNVLFHNDNRQQT 179
 176 MRLVAMLYTPLRSNGSGDDVFTVSCRVLTPTDFEFYLVPPSVESKTKFTLPIITIS 235
 180 MRLVCMLYTPLRTGGTGDSDVAVAGRVMTCPSPDENFLFLVPTVEQKTRPTFLNPLS 239
 236 ELTNSRFPPIPIBOLYAPNETNVVQNGRCRTLDGELOTTOLLSSAVCFQGRVADNG 295
 240 SLNSRAPLPISSMGISPDNVQVQNGRCRTLDGRLVGTTPVSLSHVAKRG---TSNG 296
 296 DNWDQNLQITYPNGASYDPTDEVPAPLGTQDFSGMLYGLVLTQDNNVNVSTGEAKNAGIY 355
 297 -----TVINLTLDGTPFHPF-EGPAPIGFDPDLGGCDW-----HINMTQFGHSSQTYD 344
 356 ISTTSCKFTPKIGSIGLHSHITEHVHPNOQSRT-----PVGAVDENTPFOQWVLPHYAGS 411
 345 VDTPTDFVPHLGSIOANGISGNYGVLSWISPPSHPSGSQVD-----LWKIPNYGSS 398
 412 LALNTNLAPAVATFPGEQQLFFRSRVPVCGQLQGQDAF-IDCLLPQEWNVHFYQEAAPS 470
 399 ITEATHLAPSVYPPGGEVLVFFMSKWP-----GPGAYNLPCLLPQEVISHLASEQAPT 452
 471 QADVALIRYVNPDTGRTLFEAKLHRSQFTVSHTGAY--PLVVPNGHFRFDSWVNOFYS 528
 453 VGEAALLHYVDPDTGRNLGEFKAYPDGFLTCVPGNGASSGQQPLPINGVFVFSWVSRRFYQ 512
 529 LAPMGTGNRR 539
 513 LKPVGTASSAR 523

RESULT 2
 US-09-590-020-2
 ; Sequence 2, Application US/09590020
 ; Patent No. 6355246
 ; GENERAL INFORMATION:
 ; APPLICANT: Kruger, John M
 ; APPLICANT: Maes, Roger K
 ; APPLICANT: Vilnis, Aivars
 ; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
 ; TITLE OF INVENTION: CALCIVIRUS AND VACCINES THEREOF
 ; FILE REFERENCE: MSU 4.1-446
 ; CURRENT APPLICATION NUMBER: US/09/590,020
 ; CURRENT FILING DATE: 2000-06-08
 ; PRIOR APPLICATION NUMBER: 60/138,484
 ; PRIOR FILING DATE: 1999-06-10
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 623
 ; TYPE: PR1
 ; ORGANISM: Feline calicivirus
 ; US-09-590-020-2

Query Match 9.7%; Score 278; DB 4; Length 623;
 Best Local Similarity 25.0%; Pred. No. 2.6e-18;
 Matches 129; Conservative 65; Mismatches 177; Indels 144; Gaps 26;

12 NDGAASLVPEGINETMP-----LEPVAGASIAAPVAGQTNIIIDPW-IRTNFVQAPNGEFTV 66
 81 DDGSITTPEQG---TWVGGVIAEPGSAQMSAAADMATGKSDSEWEAFFSFHTSVNW---- 133
 67 SPRNSPEILLNLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKILFAAIPP 126
 134 STSETQKILFKQSLGPLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKILFAAIPP 193

```
Db 307 NRHWSIDTPIIRPFVQANRHFDFNQETAGWSTPRFRPITITVSNNSKLIGIGVATDY 366
QY 251 TAPN-----ETNV-----VQCONGR--CTLDG-----ELQGTTLSSAVCFPL 286
Db 367 IVPGIPDGWPDITIPQLTPAGIYSITASNGDITTAAGYDAAETIVNTNFKSMYIC-- 424
QY 287 QGRTVADNGDNQDNLQIT--YPNGASYDPTDEV-----PAPLG-----TQDFSGM 331
Db 425 -GSLQRAWGDKKISNTAFITTAVRKGNSEIPSNTIDMTKLVVYQDAHVGEVQTSITILA 483
QY 332 LVGLVTQDNVNVSTGEAKNAG-----IYISTSGKTPKIGSIGLHSHITEHVHPNQOS 385
Db 484 LLOY-----TGIGEAIGSDRDKVVRISV-----LPETGARGGN-----HPIFYK 523
QY 386 RFTPVGVADVENTPFQOWVLPFYAGSLAINTNLAP 420
Db 524 NSIKLGVIRSIDVFNSQL-HTSRQLSLNNYLLP 557

RESULT 4
US-09-590-020-6
; Sequence 6, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; PRIOR FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-6

Query Match 9.6%; Score 275; DB 4; Length 622;
Best Local Similarity 24.6%; Pred. No. 5.1e-18;
Matches 126; Conservative 62; Mismatches 185; Indels 140; Gaps 24;

QY 12 NGAASLVPEGINETMP-----LEPVAGASIAAPVAGQTNIIIDPW-IRTNFVQAPNGEFTV 66
Db 81 DGSITTPEQG---TWVGGVIAEPSAQMSAAADMATGKSVDSEWEAFSFTSVNM----- 133
QY 67 SPRNSPGEILLNLELGPDLNPLYLAHLRMVNGVAGVEVQVLLAGNAFTAGKILFAAIPP 126
Db 134 STSETQKILFKQSLGPLNPLYSLAKYVVALAGSEVRFSGVFGGKLAIVVP 193
QY 127 NFLVDMISPAQITML---PHLIVDRTLEPIMPLPDVRNVFYHFNQPPRMRLVAMLY 183
Db 194 G-----IEPVQSTMLQYPHVLFDAQVEPVIFAIPDLRSNLYHLSMDTD-TTSLVIMVY 247
QY 184 TPLRNSGDDVFTVSC--RVLTPTDPDFEYILVPP-----SVESK----- 223
Db 248 NDL-INPYANDNTSSGCIIVTETKPGDFKFKLLKPPGSMITHGSPVSDLPKSSSLWTG 306
QY 224 -----TKPFTL-----PI-LTISELTNSRPPPIEQLY 250
Db 307 NRHWSIDTPIIRPFVQANRHFDFNQETAGWSTPRFRPITITVSNNSKLIGIGVATDY 366
QY 251 TAPN-----ETNV-----VQCONGR--CTLDG-----ELQGTTLSSAVCFPL 286
Db 367 IVPGIPDGWPDITIPQLTPAGIYSITASNGDITTAAGYDAAETIVNTNFKSMYICGS 426
QY 287 QGRTVADNGDNQDNLQIT-----YPNGASYDPTDEV-----PAPLG-----TQD 327
```

```
Db 427 LQRA-----WDDKKISNTAFITTAVRKGNSEIPSNTIDMTKLVVYQDAHVGEVQTS 479
QY 328 FSGMLYGLVTQDNVNVSTGEAKNAGIYISTSGKFTPKIGSIGLHSHITEHVHPNQSRF 387
Db 480 ITLALLGY-----TGIGEAIG-----SDRDKVVRISVLGETGARGGNHPIFYKNS 525
QY 388 RFTPVGVADVENTPFQOWVLPFYAGSLAINTNLAP 420
Db 526 IKLGVIRSIDVFNSQL-HSRQLSLNNYLLP 557

RESULT 5
US-09-590-020-7
; Sequence 7, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; PRIOR FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 7
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-7

Query Match 9.5%; Score 272.5; DB 4; Length 626;
Best Local Similarity 23.2%; Pred. No. 9.2e-18;
Matches 137; Conservative 84; Mismatches 211; Indels 159; Gaps 30;

QY 11 SNDGAASLVPEGINETMP-----LEPVAGASIAAPVAGQTNIIIDPW-IRTNFVQAPNGEFT 65
Db 80 ADDGSITAPEQG---TWVGGVIAEPSAQMSAAADMATGKSVDSEWEAFSFTSVNM--- 133
QY 66 VSPRNSPGEILLNLELGPDLNPLYLAHLRMVNGVAGVEVQVLLAGNAFTAGKILFAAIPP 125
Db 134 -STSETQKILFKQSLGPLNPLYSLAKYVVALAGSEVRFSGVFGGKLAIVVP 192
QY 126 NFLVDMISPAQITMLPHLIVDRTLEPIMPLPDVRNVFYHFNQPPRMRLVAMLYTP 185
Db 193 PG--VDPVOSTSMQLQYPHVLFDAQVEPVIFCLPDLRSTLYHLSMDTD-TTSLVIMVYND 249
QY 186 LRSNGSGDDVFTVSC--RVLTPTDPDFEYILVPP-----SVESK----- 223
Db 250 L-INPYANDNTSSGCIIVTETKPGDFKFKLLKPPGSMITHGSPVSDLPKSSSLWGNR 308
QY 224 -----TKPFTL-----PI-LTISELTNSRPPPIEQLYTA 252
Db 309 HWSIDTPIIRPFVQANRHFDFNQETAGWSTPRFRPISVTITEQNGAKLGIGVATDYIV 368
QY 253 PN-----ETNVQCONGRCTLDGELQGTTLSSAVCFLOGRTVADNGDNQDNLQIT 306
Db 369 PGIPDGWPDIT-----TIPGEL-----IPAGDYAITNGTGN-----IT 401
QY 307 YPNGASYDPTDEVPAPLGTQDFSGM-LYGLVTQDNVNVSTGEAKNAGIYIST-----T 359
Db 402 TATG--YTADIK---NNTNFRGMVYICSLQR-----AWGDKKISNTAFITATLDGDN 451
QY 360 SKGFTPKIGSIGLHSHIT-----EHVHPNQSR-----FTPVG---VAVDENTPFQOWV 404
Db 452 NNKINP-CNTIDQSKIYVQDAHVGVKAQTSDDTLALLGYTGIGRQAIGSDRDRVRIST 510
QY 405 LPHYAGSLAINTNLAPAVAPTFPGEQLLFFPRSPVQVQGLQODAFIDCLL---PQEWVN 461
Db 511 LPETGAR-----GGNHPIFYKNSIKLGVIRSIDVFNSQLHTSRLSLN 555
```



```

;
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/852,260
; FILING DATE: 19920619
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-B-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-852-260-2

Query Match 4.0%; Score 115; DB 1; Length 2206;
Best Local Similarity 18.7%; Pred. No. 0.23;
Matches 105; Conservative 65; Mismatches 179; Indels 212; Gaps 25;

QY 3 MASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVA-----GQTNIIDP- 50
Db 284 LAIDSMVKHNNWGIALP-----LSPDPAQDSSVEIPITVTIAPMCSFNGRLNVTAPK 338
QY 51 -----WIRTNFVQAPNG--EFTVSPR-NSPGEI----- 75
Db 339 FQGLPVLNTPGNSQVLTSDNHQSPCAIPEFDVTPIDIPGKVNMMELAEIDTMIPLNLE 398
QY 76 -----LNLLELGPDLNPLYLAH-----LSRMVNGVAGGVEVQV 107
Db 399 STKRTNMDMYRVTLSDSADLSQPIILCLSLSPAFDPRLSHTMLGEVLNYTHWAGSLKFTF 458
QY 108 LLAGNAFTAGKILFAAIPNPLVDMISPAQITMLPHLIVDVRTLPEINTPLPDVNRNVFYH 167
Db 459 LFCGSMMATGKILVAYAPPG-AQPPTSKEAMLGTHVLDLGLQSSCTMWVPWISNVTYR 517
QY 168 FNNQPO-----PRMLVAMLYTPLRNSGDDVFTVSCRVLTRPTPDFFIYL-- 215
Db 518 OTTQDSFTGGYISMFYQTRIVVPLSTPKMSMLG---FVSACN-----DFSRLLRD 567
QY 216 -----VPPSVESKTKPFTLPILTITSELNTRFPPI- 249
Db 568 TTHISQSLPQGIEDLTSEVAQALTLS-LPKQDLSLPTKASGPAHSKEVPALTAVETG 626
QY 250 YT---APNET---NVVQCQN-----GRCT---LDGE----- 271
Db 627 ATNPLAPSDTVQTRHVQRRSRSESTIESFFARGACVAIIIEVDNEQPTTTRAKLFAMWRI 686
QY 272 -LQGTTLSSAVCFLOQR-----TVADNGDNMDQNLQLTY-PNGASYDPT 316
Db 687 TYKDTVQLRRKLEFFTYSRFDMFTFVVYTNFTNANNHGHALNQ-VYQIMYIPPGA----- 740
QY 317 DEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGIYSTTSKFTPKIGSIGLHSIT 376
Db 741 ---PTPKSWDDY-----TWQTSNPSIFITYGAAPARISVYPVGLANAY 781
QY 377 EHVHPNQOSRFTFVGAVDEN 397
Db 782 SHFYDG---FAKVPLKTDAN 798

RESULT 9
US-08-461-503-2
; Sequence 2, Application US/08461503
; Patent No. 5834302
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
```

```

;
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES
; FROM CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,503
; FILING DATE: 5-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-D-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-503-2

Query Match 4.0%; Score 115; DB 2; Length 2206;
Best Local Similarity 18.7%; Pred. No. 0.23;
Matches 105; Conservative 65; Mismatches 179; Indels 212; Gaps 25;

QY 3 MASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVA-----GQTNIIDP- 50
Db 284 LAIDSMVKHNNWGIALP-----LSPDPAQDSSVEIPITVTIAPMCSFNGRLNVTAPK 338
QY 51 -----WIRTNFVQAPNG--EFTVSPR-NSPGEI----- 75
Db 339 FQGLPVLNTPGNSQVLTSDNHQSPCAIPEFDVTPIDIPGKVNMMELAEIDTMIPLNLE 398
QY 76 -----LNLLELGPDLNPLYLAH-----LSRMVNGVAGGVEVQV 107
Db 399 STKRTNMDMYRVTLSDSADLSQPIILCLSLSPAFDPRLSHTMLGEVLNYTHWAGSLKFTF 458
QY 108 LLAGNAFTAGKILFAAIPNPLVDMISPAQITMLPHLIVDVRTLPEINTPLPDVNRNVFYH 167
Db 459 LFCGSMMATGKILVAYAPPG-AQPPTSKEAMLGTHVLDLGLQSSCTMWVPWISNVTYR 517
QY 168 FNNQPO-----PRMLVAMLYTPLRNSGDDVFTVSCRVLTRPTPDFFIYL-- 215
Db 518 OTTQDSFTGGYISMFYQTRIVVPLSTPKMSMLG---FVSACN-----DFSRLLRD 567
QY 216 -----VPPSVESKTKPFTLPILTITSELNTRFPPI- 249
Db 568 TTHISQSLPQGIEDLTSEVAQALTLS-LPKQDLSLPTKASGPAHSKEVPALTAVETG 626
QY 250 YT---APNET---NVVQCQN-----GRCT---LDGE----- 271
Db 627 ATNPLAPSDTVQTRHVQRRSRSESTIESFFARGACVAIIIEVDNEQPTTTRAKLFAMWRI 686
QY 272 -LQGTTLSSAVCFLOQR-----TVADNGDNMDQNLQLTY-PNGASYDPT 316
Db 687 TYKDTVQLRRKLEFFTYSRFDMFTFVVYTNFTNANNHGHALNQ-VYQIMYIPPGA----- 740
QY 317 DEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGIYSTTSKFTPKIGSIGLHSIT 376
```

Db 741 ---PTPKSWDDY-----TWQSSNPSIFVYGAAPARISVPVYGLANAY 781
QY 377 EHVHPNQSRFTFVGAVDEN 397
Db 782 SHFYDG----FAKVPLKTDAN 798

RESULT 10
US-08-465-250-2
; Sequence 2, Application US/08465250
; Patent No. 6136570
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tacem, Joanne M.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release 1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,250
; FILING DATE: 6-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-E-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-465-250-2

Query Match 4.0%; Score 115; DB 3; Length 2206;
Best Local Similarity 18.7%; Pred. No. 0.23;
Matches 105; Conservative 65; Mismatches 179; Indels 212; Gaps 25;

QY 3 MASNDAAPSNDGAASLVEGINETMLPPLPVAGASIAAPVA-----GQTNIIIDP- 50
Db 284 LAIDSMVXHNWGAIALP-----LSPLDPAQDSSVEIPIVTIAPMCSEFNGLRNVAPK 338
QY 51 -----WIRTNFVQAPNG--EFTVSPR-NSPGEI----- 75
Db 339 FQGLPLVNTPGSNQVLTSDNHQSPCAIPEFDVTPPIDIPGEYKNNMELAEIDTWIPLNLE 398
QY 76 -----LLNLELGPDLNPLYLAH-----LSRMVNGVAGVEVQV 107
Db 399 STKRNTMDMYRVLTSDSADLSOLFCLSLSPAFDPLSHMTLGEVLNYTHWAGSLKFTF 458
QY 108 LLAGNAFTAGKILFAAIPENFLVDMISPAQITMLPHLIVDVRLEPIMTLPDVRNVEYH 167
Db 459 LFCGSMWATGKILVAVAPPG-AQPPTSRKEAMLGTHVIMDLGLQSSCTMVVFWISNVYR 517
QY 168 FNNQPO-----PRRLVAMLYTPLRSNGSGDDVFTVSCRVLTRTPDPFEIYL-- 215

Db 518 QTTQDSFTEGGYISMFYQTRIVVPLSTPKSNMLG---FVSACN-----DFSRLLRD 567
QY 216 -----VPSVESKTKPFTLPILTISELINSFPPIP-----IEQL 249
Db 568 TTHISQSALPOGIEDLTSEVAGGALTLS-LPKQDSLPDPTKASGPAHSKEVPALTAVETG 626
QY 250 YT---APNET---NVVQCON-----GRCT---LDGE----- 271
Db 627 ATNPLAPSDTVQTRHVQVRRSRSESTIESFFARGACVAIIIEVDNEQPTTTRAKLFAWMRI 686
QY 272 -LQGTTLQLSSAVCFLOOR-----TVADNGDNWDLQLQTY-PNGASYDPT 316
Db 687 TYKDTVQLARKLEFFETYSRDFMEFTFVVTAFTNANNHALNQ-VYQIMYIPPGA----- 740
QY 317 DEVPAPLGTQDFSGMLYGVLTQDNVNVSTGBAKNAKGIYISTTSKFTPKIGSILHSIT 376
Db 741 ---PTPKSWDDY-----TWQSSNPSIFVYGAAPARISVPVYGLANAY 781
QY 377 EHVHPNQSRFTFVGAVDEN 397
Db 782 SHFYDG----FAKVPLKTDAN 798

RESULT 11
5516630-2
; Patent No. 5516630
; APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; FEINSTONE,
; STEPHEN M.; PURCELL, ROBERT H.; RACANIELLO, VINCENT R.;
; BAROUDY, BAHIGE M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/788,262
; FILING DATE: 06-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 256,135
; FILING DATE: 06-OCT-1988
; APPLICATION NUMBER: 654,942
; FILING DATE: 27-SEP-1984
; APPLICATION NUMBER: 537,911
; FILING DATE: 30-SEP-1983
; SEQ ID NO:2:
; LENGTH: 1091
; 5516630-2

Query Match 3.8%; Score 108.5; DB 6; Length 1091;
Best Local Similarity 21.6%; Pred. No. 0.31; 136; Indels 133; Gaps 22;
Matches 87; Conservative 46; Mismatches 136; Indels 133; Gaps 22;

QY 96 YNGVAG-GVEVQVLLAGNAFTAGKILFAAIPPNFLVDMISPAQITMLPH-LIVDVRLE 152
Db 360 YHTYARFGIEIQVINPTFPQGGGLICAMVPGDQSYGSI--ASLTVPYHGLLNCNINNV 417
QY 153 PIMTLPDVRNVYFHNNOPOPRMRLVAMLYTPLRSNGSGDDVFTVSCRVLTRTPDPDEF 212
Db 418 RIKVFFIYTRGA-YHKFDQYVPMELTIRVWSEL-NIGTGTSAYT-SUNVLARFT-DUEL 473
QY 213 IIVLPSPVESKTKP-----TLPILTISELTNSR----- 241
Db 474 HGLTLPSTOMNEFRVSTTENNVNLSNVEBARAKMSFALQOEDWKSQSGGKIKITHF 533
QY 242 -----FP-----IPIEQLYTAPNETNVVQCONCRCTLDGELQTT 276
Db 534 TTWTSIPTLAAQFPFNASDSVQGIKIVDPVDFPQMTNTNPDQ---KC-----IT 581
QY 277 QLLS--SAVCFIQGRTVADNGDNWDLQLQTYPGA---SYDPTDEV----- 319
Db 582 ALASIQMFCFWRGDLV-----FDQVPTTKYHSGRLLFCFVPGNELIDVSGITLKQAT 635
QY 320 PAPLGTQDFSGMLYGVLTQDNVNVSTGBAKNAKGIYISTTSKFTPKIGSILHSIT 363
Db 636 TAPCAVMDITGVQSTLRFRVPMISDTPYRVNRYTKSA-HOKGEY--TAIGKLVVYCNRL 692

QY 364 -TPKIGSIGLHSITHEVHPN-----QQRFTPVGVAVDENT 398
Db 693 TSPS-----NVASHVRVNVVLSAINLECFAPLVHAMDVT 727

RESULT 12

US-08-475-886-2

; Sequence 2, Application US/08475886A

; Patent No. 6113912

; GENERAL INFORMATION:

; APPLICANT: FUNKHOUSER, ANN W

; APPLICANT: EMERSON, SUZANNE U

; APPLICANT: PURCELL, ROBERT H

; APPLICANT: D'HONDT, ERIC

; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

; FILE REFERENCE: 20264262US2

; CURRENT APPLICATION NUMBER: US/08/475,886A

; CURRENT FILING DATE: 1995-06-07

; EARLIER APPLICATION NUMBER: 07/947,338

; EARLIER FILING DATE: 1992-09-18

; EARLIER APPLICATION NUMBER: 08/397,232

; EARLIER FILING DATE: 1995-03-10

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 2

; LENGTH: 2227

; TYPE: PRT

; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175

US-08-475-886-2

Query Match 3.8%; Score 108.5; DB 3; Length 2227;
Best Local Similarity 21.6%; Pred. No. 1;
Matches 87; Conservative 46; Mismatches 136; Indels 133; Gaps 22;

QY 96 YNGYAG-GVEQVQLAGNAFTAGKILFAAIPNPLVDMISPAQITMLPH--LIVDVRTLE 152

Db 123 YHTVARFGIEIQVINPTPFQGGGLICAMVPGDQSYGSI--ASLTVYPHGLINCINN 180

QY 153 PIMTLPDVRNVFYHFNQPOPRMELVAMLYTPLRNSGDDVFTVSCRVLTRPTPDEF 212

Db 181 RIKVPFIYTRGA-YHFKDPQYPVWELTIRVWSEL-NIGTGSAYT-SLNVLARFT-DLEL 236

QY 213 IYLVPPSVESKTKPF-----TLPILTISELTNSR----- 241

Db 237 HGLTPLSTQMRNEFRVSTTENVNLSNYEDARAKMSFALDQEDWKSDPSQGGIKITHF 296

QY 242 -----FP-----IPISOLYAPNETNVVQCGRCITLDGELQGT 276

Db 297 TTWTSIPTLAAQFPFNASDSVGGQIKVIPDPYFFQMTNTNP 344

QY 277 QLLS--SAVCFQGRVTADNGDNWQNLQLTYPNGA---SYDPTDEV----- 319

Db 345 ALASIQCMFCFWRGLV-----FDQVFPPTYKHSGRLLFCFVPGNELIDVSGITLKQAT 398

QY 320 PAPLGTQDFSGMLYGVL-----TQDNVNVSTGEAKNAKGIYISTTSGRF----- 363

Db 399 TAPCAVMDITGVQSTLRFRVPMISDTPYRVNRYTKSA-HQGEY--TAIGKLIVVCYNRL 455

QY 364 -TPKIGSIGLHSITHEVHPN-----QQRFTPVGVAVDENT 398

Db 456 TSPS-----NVASHVRVNVVLSAINLECFAPLVHAMDVT 490

RESULT 13

US-08-475-886-4

; Sequence 4, Application US/08475886A

; Patent No. 6113912

; GENERAL INFORMATION:

; APPLICANT: FUNKHOUSER, ANN W

; APPLICANT: EMERSON, SUZANNE U

; APPLICANT: PURCELL, ROBERT H

; APPLICANT: D'HONDT, ERIC

; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

Wed Jun 2 09:13:34 2004

Query Match 3.8%; Score 108.5; DB 3; Length 2227;
Best Local Similarity 21.6%; Pred. No. 1;
Matches 87; Conservative 46; Mismatches 136; Indels 133; Gaps 22;
QY 96 YNGYAG-GVEYQVLLAGNAFTAGKILFAAIPNPNFLVDMISPAQITMLPH--LIVDVRTLE 152
DB 123 YHTYARFGIEIQVQINPTFPQGGGLICAMVFGDQSYGSI--ASLTVPYHGLLNCNINNV 180
QY 153 PIMTLPDVRNVYHFNQOPQPMRLVAMLYTPLRNSGDDVFTVSCRVLTPTDPEF 212
DB 181 RIKVPFIYTRGA-YHFKDQYVPMELTIRVWSEL-NIGTGTSAYT-SLNVLARFT-DLEL 236
QY 213 IYLVPPSVESKTKPF-----TLPILTISELTNSR----- 241
DB 237 HGLTLPSTQMRNEFRVSTTENVNLSNVEDARAKMSFALDOEDWKSDPSQGGGIKITHF 296
QY 242 -----FP-----IPIEQLYTAPNETNVVQCQNGRCTLDGELQGT 276
DB 297 TWTISPTLAAQFPFNASDSVGQIQKIVDPYFFQMTNTPDQ-----KC-----IT 344
QY 277 QLLS--SAVCFLOGRTVADNGNDQNLQLTYPNGA--SYDPTDEV----- 319
DB 345 ALASICQMFCEWRGDLV-----FDQVFPPTKTHSGRLLFCFVPGNELIDVSGITLKQAT 398
QY 320 PAPLGTQDFSGMLYGVL-----TQDNVNVSTGEAKNAKGIYISTTSKGF----- 363
DB 399 TAPCAVMDITGVOSTLRFVRPWIISDTPYRVNRYTKSA-HQKGEY--TAIGKLIIVCYNRL 455
QY 364 -TPKIGSIGLHSITEHVHPN-----QOSRFTPVGVAVDENT 398
DB 456 TSFS-----NVASHVRNVVYLSAINLECFAPLYHAMDVTT 490

RESULT 15
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262051
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2
Query Match 3.8%; Score 108.5; DB 3; Length 2227;
Best Local Similarity 21.6%; Pred. No. 1;
Matches 87; Conservative 46; Mismatches 136; Indels 133; Gaps 22;
QY 96 YNGYAG-GVEYQVLLAGNAFTAGKILFAAIPNPNFLVDMISPAQITMLPH--LIVDVRTLE 152
DB 123 YHTYARFGIEIQVQINPTFPQGGGLICAMVFGDQSYGSI--ASLTVPYHGLLNCNINNV 180
QY 153 PIMTLPDVRNVYHFNQOPQPMRLVAMLYTPLRNSGDDVFTVSCRVLTPTDPEF 212
DB 181 RIKVPFIYTRGA-YHFKDQYVPMELTIRVWSEL-NIGTGTSAYT-SLNVLARFT-DLEL 236
QY 213 IYLVPPSVESKTKPF-----TLPILTISELTNSR----- 241
DB 237 HGLTLPSTQMRNEFRVSTTENVNLSNVEDARAKMSFALDOEDWKSDPSQGGGIKITHF 296

QY 242 -----FP-----IPIEQLYTAPNETNVVQCQNGRCTLDGELQGT 276
DB 297 TWTISPTLAAQFPFNASDSVGQIQKIVDPYFFQMTNTPDQ-----KC-----IT 344
QY 277 QLLS--SAVCFLOGRTVADNGNDQNLQLTYPNGA--SYDPTDEV----- 319
DB 345 ALASICQMFCEWRGDLV-----FDQVFPPTKTHSGRLLFCFVPGNELIDVSGITLKQAT 398
QY 320 PAPLGTQDFSGMLYGVL-----TQDNVNVSTGEAKNAKGIYISTTSKGF----- 363
DB 399 TAPCAVMDITGVOSTLRFVRPWIISDTPYRVNRYTKSA-HQKGEY--TAIGKLIIVCYNRL 455
QY 364 -TPKIGSIGLHSITEHVHPN-----QOSRFTPVGVAVDENT 398
DB 456 TSFS-----NVASHVRNVVYLSAINLECFAPLYHAMDVTT 490
Search completed: June 1, 2004, 13:58:06
Job time : 14.9134 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 34.648 Seconds
(without alignments)
4368.312 Million cell updates/sec

Title: US-09-926-799-9
Perfect score: 2863
Sequence: 1 MKMASNDAPSDGASLVP.....VNQFSLAPMTGNGRRRIQ 542

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1179	41.2	530	US-10-314-739-3	Sequence 3, Appli
2	276.5	9.7	547	US-10-670-695-4	Sequence 4, Appli
3	276.5	9.7	671	US-10-670-695-2	Sequence 2, Appli
4	265.5	9.3	668	US-10-209-507-4	Sequence 4, Appli
5	265.5	9.3	669	US-10-209-507-2	Sequence 2, Appli
6	145.5	5.1	1147	US-10-327-481A-38	Sequence 38, Appli
7	133	4.6	6310	US-10-282-122A-67793	Sequence 67793, A
8	116	4.1	394	US-10-104-047-3560	Sequence 3560, Ap
9	114.5	4.0	594	US-10-282-122A-65399	Sequence 65399, A
10	112.5	3.9	3930	US-10-282-122A-46817	Sequence 46817, A
11	111.5	3.9	3317	US-10-080-334-200	Sequence 200, App
12	110.5	3.9	5376	US-10-028-248A-74	Sequence 74, Appli
13	110.5	3.9	5376	US-10-107-782-74	Sequence 74, Appli
14	110.5	3.9	6304	US-10-147-026-16	Sequence 16, Appli
15	109.5	3.8	1109	US-10-389-566-1196	Sequence 1196, Ap

16	109.5	3.8	1109	15	US-10-389-566-2378	Sequence 2378, Ap
17	109.5	3.8	2468	12	US-10-282-122A-66335	Sequence 66335, A
18	109.5	3.8	2468	14	US-10-246-330-4	Sequence 4, Appli
19	108.5	3.8	836	14	US-10-272-459-40	Sequence 40, Appli
20	108.5	3.8	980	14	US-10-272-459-41	Sequence 41, Appli
21	108.5	3.8	2227	9	US-09-929-955-12	Sequence 12, Appli
22	108.5	3.8	2227	13	US-10-104-366-12	Sequence 12, Appli
23	108.5	3.8	2227	13	US-10-135-988-2	Sequence 2, Appli
24	108.5	3.8	2227	13	US-10-135-988-4	Sequence 4, Appli
25	108.5	3.8	2227	13	US-10-135-988-6	Sequence 6, Appli
26	108.5	3.8	2227	16	US-10-719-619-12	Sequence 12, Appli
27	107.5	3.8	2777	12	US-10-423-483-4	Sequence 4, Appli
28	107.5	3.8	2777	15	US-10-220-587-4	Sequence 4, Appli
29	107.5	3.8	3399	15	US-10-080-334-196	Sequence 196, App
30	105.5	3.7	368	14	US-10-156-761-7936	Sequence 7936, Ap
31	105	3.7	547	15	US-10-080-334-276	Sequence 276, App
32	105	3.7	596	12	US-10-188-248-68	Sequence 68, Appli
33	104.5	3.7	1194	12	US-10-282-122A-46577	Sequence 46577, A
34	104	3.6	925	15	US-10-108-260A-3889	Sequence 3889, Ap
35	103	3.6	624	15	US-10-258-106-4	Sequence 4, Appli
36	103	3.6	913	15	US-10-369-493-2294	Sequence 2294, Ap
37	102.5	3.6	698	15	US-10-074-978A-144	Sequence 144, App
38	101.5	3.5	390	11	US-09-764-875-675	Sequence 675, App
39	101.5	3.5	1469	12	US-10-262-511-164	Sequence 164, App
40	101.5	3.5	2753	12	US-10-262-511-160	Sequence 160, App
41	101.5	3.5	2753	12	US-10-262-511-166	Sequence 166, App
42	101.5	3.5	3838	12	US-10-262-511-162	Sequence 162, App
43	101.5	3.5	5374	15	US-10-028-248A-75	Sequence 75, Appli
44	101.5	3.5	5374	15	US-10-107-782-75	Sequence 75, Appli
45	101	3.5	1004	12	US-10-114-270-32	Sequence 32, Appli

ALIGNMENTS

RESULT 1
US-10-314-739-3
; Sequence 3, Application US/10314739
; Publication No. US20030129588A1
; GENERAL INFORMATION:
; APPLICANT: Bates, Mary K
; Jiang, Xi
; Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; Characterize No. US20030129588A1walk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pubbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Ave., N.W.
; CITY: Washington, D.C.
; STATE: <Unknown>
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/10/314,739
; APPLICATION NUMBER: US/10/314,739
; FILING DATE: 09-Dec-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,049
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311.023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-314-739-3

Query Match 41.2%; Score 1179; DB 14; Length 530;
Best Local Similarity 46.5%; Pred. No. 2.2e-102;
Matches 256; Conservative 84; Mismatches 171; Indels 40; Gaps 12;

QY 1 MOWASNDAPSNDGAA---SLVPEGINETMPL--EPVAGASIAAPVAGQTNIIIDPWIRTN 55
DB 1 MMWASKDATSVDGAGAGQLVPE-VNASDPLAMPVAGSSAVATAGQVNPIDPWIINN 59

QY 56 FVQAPNGEFTVSPRNSPGSEILLNLELGPDLNPLYLAHLRMYNGYAGGVEQVLLAGNAFT 115
DB 60 FVQAPQGETISPNNTPGDVLFDLSLGPLNPLFLLHLSQMYNGWYGNRMVRIMLAGNAFT 119

QY 116 AGKILFAALPPHFLMDISPAQITMLPHLIVDRTLEPIMTLPDVRNVFVHFNNOQOPR 175
DB 120 AGKIIIVSCIPGCGSHNLIAQAATLPHVIADRTLDPIEVLDPVRNVLFHNDENQOT 179

QY 176 MRLVAMLYPLRNSGGDDVFTVSCRVLTRTPDFEFYLVPPSVESKTKPTLPILTIS 235
DB 180 MRLVCMLYPLRTGGGTGDSFVAGRVMTCPSPDFNFLVLPVETVEQKTRPFTLPNPLS 239

QY 236 ELTNSRPPPIEQLYTAPNETNVVQNGRCTLGDELQTTOLLSAVCFLOGRVADNG 295
DB 240 SLNSRAPLPISSMIGSPDNVQSVQFQNGRCTLGDLGLVGTTPVLSHVAKIRG---TSG 296

QY 296 DNDQNLLOLTYPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGY 355
DB 297 ---TVINUTELDTGTFPHF-EGPAPIGPFDLGGCDW-----HNMTOFGHSSQTYD 344

QY 356 ISTSGKFTPKIGSIGLHSITEHVHPNQQRFT-----PVGAVDENTPQOVLPHYAGS 411
DB 345 VDTTPTDFVPHLGSIQANGIGSGNYGVLSWISPPSHPSGSQVD-----LWKIPNYGSS 398

QY 412 LALNTNLAPAVAPTFPGEQLLFFSRVPCVQGLQGQDAF-IDCLLPQEWYNNHFYQAAAPS 470
DB 399 ITEATHLAPSVPYPGGEVLVFFMSKP-----GPCAYNLPCLLPQEYISHLASEQAPT 452

QY 471 QADVALIRYNPDGRTLFEAKLHRSGETTVSHTGAY--PLVVPNGHFRFDSWYVQFYS 528
DB 453 VGEAALLHYVDPTGRNLGEPKAYPDGFLTCVPNGASSGPPQLPFGVVFVSVGRFYQ 512

QY 529 LAPMGTGNRR 539
DB 513 LKPVGTASSAR 523

RESULT 2
US-10-670-695-4
; Sequence 4, Application US/10670695
; Publication NO. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT FILING DATE: 2003-09-25
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4

LENGTH: 547
TYPE: PRT
ORGANISM: Feline calicivirus
US-10-670-695-4

Query Match 9.7%; Score 276.5; DB 12; Length 547;
Best Local Similarity 25.1%; Pred. No. 6.7e-17;
Matches 131; Conservative 68; Mismatches 171; Indels 151; Gaps 28;

QY 11 SNDGAASLVPEGINETMP-----LEPVAGASTAAPVAGQTNIIIDPW-IRTNFVQANGGEFT 65
DB 1 ADDGSITAEQEG---TMVGVIAPESQAQMTAADWATKSDVSEWEAEFSEHTSVNW--- 54

QY 66 VSPRNSPGSEILLNLELGPDLNPLYLAHLRMYNGYAGGVEQVLLAGNAFTAKILFAAIP 125
DB 55 -STSETQGGKILFKQSLGPLNPLYLAHLAKLYVAMSSIEVRFSISGSGVFGKLAIVVP 113

QY 126 PNFVDMISPAQITMLPHLIVDRTLEPIMTLPDVRNVFVHFNNOQOPRMLVAMLYTP 185
DB 114 PG--VDPVQSTSMQYPHVLFDARQVPERVIFCLPDLRSTLYHLMSTD-TTSLVIMVND 170

QY 186 LRNGSGDDVFTVSC--RVLTRTPDFEFYLVPP-----SVES---KT----- 224
DB 171 L-INPYANDANSSGCIIVTETKPGDFKPHLLKPPGSMLTGHSIPSDLIPTKSSLMIGNR 229

QY 225 -----KPFTL-----PI-LTISELTNSRPPPIEQLYTA 252
DB 230 YWSDITDFVIRPFVFOANRHFDFNQETAGWSTPRFRPISVTITEQNGAKLGIGVATDYIV 289

QY 253 PN-----ETNVQCCNGRCTLGDEL-----OGTTOLLSAVCFLOGRVAD---N 294
DB 290 PGIPDGWPD-----TIPGELIPAGDYAITNGTGNDITAT-----GYTADIILKN 335

QY 295 GDN-----W-DQNLQLTYPNGASYD-PTDEVPAPLGTQDFSGMLYGVLTQD 339
DB 336 NTNFRGMVIGSLQORANGDKKISNTAFITATLDGDNKKINPCNTIDQSKI---VVFQD 392

QY 340 N-----VNVS-----TGEAKNAKG-----IYISTSGKFTPKIGSIGLHSITEHV 379
DB 393 NHVGKKAQTSDDTLALLGYTIGGEQAIGSDRDRVRVRIST-----LPETGARGN----- 441

QY 380 HPNQSRFTPVGVAVDENTPQOVLPHYAGSLALNTNLAP 420
DB 442 HPFYNSIKLYVIRSIDVENSQIL-HTSRQLSLNHLVLLP 481

RESULT 3
US-10-670-695-2
; Sequence 2, Application US/10670695
; Publication NO. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT FILING DATE: 2003-09-25
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 671
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-670-695-2

Query Match 9.7%; Score 276.5; DB 12; Length 671;
Best Local Similarity 25.1%; Pred. No. 9.3e-17;
Matches 131; Conservative 68; Mismatches 171; Indels 151; Gaps 28;

Qy	11	SNDGAASLVPGINBTMP-----LEPVAGASTAAPVAGQTNIIDPW-IRTNFVOAPNGEPT	65
Db	125	ADDSGITAPEQG---TWGGVIAEPSQAMSTAADMATGKSDVSEWAEFPFSTSVNW---	178
Qy	66	VSPRNSPGEIILNLELGPDLNPNYLAHLSRMVNGYAGGVEQVLLAGNAFTACKILFAAP	125
Db	179	-STSETQCKILFKOSGLPPLNPNYLBELHAKLYVWMSGSEIVRFSISGSGVFGKLAIVVP	237
Qy	126	PNFLVDMISPAQITMLPHLIVDVRLEPIMTLPDPVRNVFYHFNQPOPRMELVAMLYTP	185
Db	238	PG--VDPVQSTSMLOYPHVLFDARQVEPIFCLPDLRSTLYHLSMDTD--TTSIVIMVYND	294
Qy	186	LRNSGSGDDVETVSC--RVLTRPTDPDRFYLYVPP-----SVGS-----KT-----	224
Db	295	L-INPYANDANSSGCVIVETKPGDPFKHLLKPPGSMLTGHSIPSDLIPTKTSLSMIGNR	353
Qy	225	-----KPFTL-----PI-LTISELTNSRFPPIEQLYTA	252
Db	354	YWSDITDEVIAPFPVQANRHFDNFQETAGWSTPRFRPISVLTTEONGAKLGVATDYIV	413
Qy	253	PN-----ETNVQOQNGRCTLDGEL-----QGTQLLSSAVCFLQGRVAD---N	294
Db	414	PGIPDGWPDT-----TIPGELIPAGDYAITNGTGNDDTTAT---GYDTADIKN	459
Qy	295	GDN-----W-DQNLLOLYTPNGASYD-PTDEVPAPLGTQDPFGSMLYGVLTOD	339
Db	460	NTNFRGMVYICGSLQRAWGDKKISNTAFITTATLDGDNNNKINPCNTIDQSKT---VWFOD	516
Qy	340	N-----VNVS-----TGEAKWAG-----IYISTSGKFTPKGISGLHSITEHV	379
Db	517	NHVCKAKQTSDDTLALLGYTGIGEQAIGSDRDRVRIST-----IPEFGARGN-----	565
Qy	380	HPNQCSRFTPVGVAVDENTFPQQWVLPHYAGSLALNTLAP	420
Db	566	HPIEYKNSIKLGYVIRSIDFVNSOIL-HTSROLISLNHYLLP	605

RESULT 4
 US-10-209-507-4
 ; Sequence 4, Application US/10209507
 ; Publication No. US20030105033A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Audonnet, et al.
 ; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V
 ; FILE REFERENCE: 454313-3151.2
 ; CURRENT APPLICATION NUMBER: US/10/209,507
 ; CURRENT FILING DATE: 2002-11-12
 ; PRIOR APPLICATION NUMBER: 09/617,594
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: 60/193,332
 ; PRIOR FILING DATE: 2000-03-30
 ; PRIOR APPLICATION NUMBER: France 00 01761
 ; PRIOR FILING DATE: 2000-02-11
 ; PRIOR APPLICATION NUMBER: France 99 09421
 ; PRIOR FILING DATE: 1999-07-16
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 668
 ; TYPE: PRT
 ; ORGANISM: Feline calicivirus
 ; US-10-209-507-4

```

Db      181  SETQKILFKQSLGFLNPNYLTHLAKLVAVWSGSIEVRFSGSGVFGKLAIVVPPG- 239
Qy      129  LVDMSPAQITMLPHLIVDVRTLEPIIMTPLPDVRNVFYHFNNQPPQRMRLVAMLYTPL-- 186
Db      240  -IDPVQSTSMLOYPHVLFDARQVPEVFTIEDLRNSLTHLMSDTD-TTSLVIMIYNDLIN 297
Qy      187  ----RNSGDDVFTVSCRVLTRTPDPEFIYLVPPSVESKTKPFTLPILITISELTSNRP 242
Db      298  PYANDSNSGCIV-----TVETKPGDFKPHLLKPPG-----SMLTHGSI 337
Qy      243  P 243
Db      338  P 338

RESULT 5
US-10-209-507-2
; Sequence 2, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RES
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-2

```

```

; PUBLICATION NO.: US20030109303A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 03/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
; US-10-209-507-4

```

Wed Jun 2 09:13:34 2004

```

Db      500 ENDTLL---VSASAGKO-FELRL 518

RESULT 7
US-10-282-122A-67793
; Sequence 67793, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zvekund, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67793
; LENGTH: 6310
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-10-282-122A-67793

Query Match      4.6%; Score 133; DB 12; Length 6310;
Best Local Similarity 23.2%; Pred. No. 0.13;
Matches 113; Conservative 58; Mismatches 201; Indels 116; Gaps 26;

QY      7 DAAPS-----NDGASLNPPEGINETMPLPVPAGASIAAPVAGQTNIIDPWIRTNFVQAPN 61
Db      3575 NASPSATFTADSTAPAPAGLTITPGGSAIOGSEAGSTV-EVKLADGTLGVTVVPAG 3633
QY      62 GEFTV--SPRNSPGEILLNLEL-----GPDLPYLAHLRMYNGYAGGV 105
Db      3634 GSTVPLSPAQLDQG-ALNVTLTDAAGNISQPSQIFAPDITP-----PALPTDV 3681
QY      106 QVLLAGNAFT-----AGKILFAALPP--NELVDMISPAQ-----ITMLPHLI 145
Db      3682 AVSSDGTAVTGNAPGASSVTVSDGAGNVITVAVNPDGSESVPLDTFQNNQQTTVV---VV 3738
QY      146 VDV--RTLEPIMTLPDVRNVFYHFNQPPQPRNRLVAMLYTPLRS--NCGDDVFTVSCR 201
Db      3739 TDAAGNDSAPFVSVTAPDTTN-----BEPATGLTV---SPDGSVTGGTAEPGSTVEVR 3787

US-10-327-481A-38
; Sequence 38, Application US/10327481A
; Publication No. US20040001864A1
; GENERAL INFORMATION:
; APPLICANT: King, Andrew M.O.
; APPLICANT: Burman, Alison J.
; APPLICANT: Audonnet, Jean-Christophe F.
; APPLICANT: Lombard, Michel F.A.
; TITLE OF INVENTION: Vaccine against Foot-and-Mouth Disease
; FILE REFERENCE: 454313-3178
; CURRENT APPLICATION NUMBER: US/10/327,481A
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: PCT/FR01/02042
; PRIOR FILING DATE: 2001-06-27 0008437
; PRIOR APPLICATION NUMBER: FR 0008437
; PRIOR FILING DATE: 2000-06-29
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 1147
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
US-10-327-481A-38

Query Match      5.1%; Score 145.5; DB 15; Length 1147;
Best Local Similarity 20.8%; Pred. No. 0.00056;
Matches 92; Conservative 47; Mismatches 203; Indels 101; Gaps 13;

QY      70 NSPGEILLNLELGPDLNLYLAHLSRMVNYAGGVEQVLLAGNAFTAGKILFAAIPNPL 129
Db      158 DKPFGYLTKEJLPTDHHGVGHVGLVDSYAYMRNGMDVEVSAGVNGFNGGCLLVAMVPEWKA 217
QY      130 VDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNQPPQPRNRLVAMLYTPLRSN 189
Db      218 FDTREKYQUTLFFHQFI-----SPRTNMTAHTVYPLGV 251
QY      190 GSGDDVFTVSCRVLTPTDPEFFIYLVPPSVESKTPFTLPILITISELTNSRFPPIEQOL 249
Db      252 NRYDQY-----KQHKPMTLVVMVLSPLTVSNTAAPQIKV 285
QY      250 YT--APNETNVQCGNCRCTLDGELQGTTLQLSSAVCFELQRTVADNGDNDQNLLOQTY 307
Db      286 YANIAPTYVHVA-----GELPSKEGIFPVACADYGGLVTTDPKTAADPVYGVKN 335
QY      308 PNGASY-----DPTDEVPAPLGTQDFSGMLYGVLTQDNV-----NVSTGEAKNAKG 353
Db      336 PPKTNPYGRFTNLLDVAEACPTFLREDD--GKPYVTVTRADDTRLLAKFDVSLA-AKHSN 392
QY      354 IYISTSGKFTPKIGSIGHLSI---TEHVHPNQSRFTVGVAVDENTPFQOWVLPHYAG 410
Db      393 TYLSGIAQYQYQTSYSGTINLHFMFTGTDSKARYWVAYIPPGVETPDPPEEAHCIAEW 452
QY      411 SLALNTNLAPAVAPTFPGEQLLFFRSRVCVQGLQGDADFICLLLPQEWNVHIFYEAPRS 470
Db      453 DTGLNSKFTESI--PYVSAADYAYTASDTAETTNVQ-----WVCVYQTHGKA 499
QY      471 QADVALIRYNPDGRTLFEAKL 493
```

QY 202 ---VLTRPT---PDEFTYLVPPSVESKTKFTTILITISLTNSRPIPIEQIYTAPN 254
Db 3788 NPDDTVRGVTVAGPDGTVIPAPPLASGE---TVDVVVIDPAGNESPEIPL---TGPT 3840
QY 255 ETNVVQCQNGRCTLGGEL---CGTTQLL---SSAVCFLOQRTVADNG-----DNWDONL 302
Db 3841 GTEVATPSALATSVGFLLTGGTGVSLITVTSGGTTLGSATVSGDGTFRVFQNAQLNA 3900
QY 303 LQITVPNGASYDPTDEVAPLGTQDFSGMLYGVLTQDNVNVN-----TGAKNAKGIYIST 358
Db 3901 QILQVSATVGDQPSVPAIIIVANDTTAP--DAPTQVNLNATGSLTGTGQEVGATVRVTD 3958
QY 359 TSGKF--TPKIGSIGLHSIT-----EYHVPNQSRFTPVGVAVDENTPFOQWVLPY 408
Db 3959 LQGTLLGTATVDSNGLSVSFSFPAVANGONLIVTQADAAGNVSLAATLQAPDLQ--APLA 4016
QY 409 AGSLALNT 416
Db 4017 ASNLSINS 4024

RESULT 8

US-10-104-047-3560
; Sequence 3560, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3560
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3560

Query Match 4.1%; Score 116; DB 15; Length 394;
Best Local Similarity 21.8%; Pred. No. 0.063;
Matches 74; Conservative 49; Mismatches 103; Indels 114; Gaps 17;
QY 106 QVLLAGNAFTACKILFAAIPPNFLVDMISPAQITMLPHLIVDVRLTLEPMTPLP----- 159
Db 62 RVFLIGTS-TIG-----PPN-----PVSCATPSLLSLVLKPCPLESLILPLGRKCTY 109
QY 160 DVNVVYHFNNQPPRMRLVAMLYTPLRSNGSGDDVFTVSCRVLTRPTDPFEFYLVPSS 219
Db 110 GIKCRFFH-PERPSCPQSVSA---DELRAANA-----LSPPR 142
QY 220 VESKTKPFTLITILITISLTNSRPIPIEQIYTAPNETNVVQCQNGRCTLDGE----- 271
Db 143 APSKDK-----NGRRPSSQ-----SSLLTSESCSLDGKGLGAQASP 182
QY 272 ---LQGTQLLSAVCFLOQRTVADNGDNWDONLQLITYPNGASYDPTDEVAPLGTQDF 328
Db 183 GSRQGLTQTYAPS-----GRSLAPSGG-----SSSGPTDMLPQTLDLSLPY 225
QY 329 SGMLYGVLTQDNVNVSTGEAKNAKGIYISTTSGKFTPKIGSIGLHSITEHVHPN--QQRFF 387
Db 226 -----VSQCLDSDGIGSLE-----SQMSLWGRVGGPG-----EPGPPRPIYTCY 266
QY 388 TPVGAVDENTPFOQWVLPYHAGSALNTNLA PAVATPP 427
Db 267 SPYGSLEPATAFAFGAMGAGHFSVPADYPPA--PPAPP 305

RESULT 9

US-10-282-122A-65399
; Sequence 65399, Application US/10282122A

Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 65399
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-282-122A-65399

Query Match 4.0%; Score 114.5; DB 12; Length 594;
Best Local Similarity 19.9%; Pred. No. 0.17;
Matches 92; Conservative 56; Mismatches 174; Indels 141; Gaps 20;
QY 44 QTNIIDPWIRTNFVQAPNGEFTVSPRNSPGBILLNLGLPDLNPLYLAHLRMYN---GY 99
Db 19 QTVTVKGVHRR--RDHGGVIFIDLRDREG--IVQVVIDPDTPEAFAAADSARNEYVLGI 74
QY 100 AGGVEVQVLLAGNAFTACKI-----LFAAIPPNFLVDMISPAQITMLPHLIVDVRT 150
Db 75 AGRVNRPKTTNDKMSCKIEILAKEIEVLNAAATFPFQIDDENISENVRLNTRVIDLR- 133
QY 151 LEPIMTLPDVRNVYHFNNQPPRMRLVAMLYTPLRSNGSGDDVFTVSCRVLTRPTDPF 210
Db 134 -RPVN-----QRNLRURYQVATGVRYLDQAQGFIDISTPLMTRSTPEG 175
QY 211 EFIVLVPPSVSKTKP---FTLP-----ILTIS-----ELTNSRPP- 243
Db 176 ARDYLVP-----SRVHPGEFFALPOSQFLKOLLWVAGFDRIYQITKCFRDEDLRADROPE 231
QY 244 ---IPIEQIYTAPNETNVVQCQNG-----RCTLDGELQGTQLLSAVCFLOQRTVAD- 293
Db 232 FTQIDLETSLF--NEDEIMDITEGMAKQVFKDALNVLDGDFRPMYSEAMFYGSKPDM 289
QY 294 -----NGDNWDONLQLITYPNGASYD--PTDEVAPLGTQDF 328

Db 290 RINLKFTLTLMTKEEFKVFPGAADMKGRVVALRVNPNVNGAKSRKEIDRYTKFVGIYGA 349
Qy 329 SMLXGVLTDQNVNVTGE-----AKNAKGIYISTTSKGF 363
Db 350 KGLAY-IKVNDAGNLSNEDSGLOSPIVKFLSENALKEIERTAAQNGDIIFFGADKAKY 408
Qy 364 TPK-IGSGLHSITHVHPNQSRFTFVGVAVDENTPQOAVL 405
Db 409 VNEAVGALKRIKVGLEHGKDN-----GYFTDEWKPL--WV 441

RESULT 10
US-10-282-122A-46817
; Sequence 46817, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangou
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46817
; LENGTH: 3930
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (19)..(19)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (38)..(38)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (3455)..(3455)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE

; LOCATION: (3459)..(3459)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (3552)..(3552)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (3566)..(3566)
; OTHER INFORMATION: X=any amino acid
; US-10-282-122A-46817
Query Match 3.9%; Score 112.5; DB 12; Length 3930;
Best Local Similarity 18.9%; Pred. No. 5.4;
Matches 136; Conservative 91; Mismatches 262; Indels 229; Gaps 34;
Qy 9 APSNDGAAS-----LVPEG---INETMPL-----EPVAGASTAAPVAGOT---- 45
Db 2203 ALTNNNGSSAQNVIPTDTVPSTAFIADTFSTINGIPQSGAMPVNGVNGISITAGTTTVS 2262
Qy 46 -----NLIIDPWIRTF--VOAPNGEFTVSPRNSP--GEILLNLELGPDLN---- 86
Db 2263 FQVTVTSLPTENPIVNFSSISYQLVSPDAETISINPVSTQIKAILSMKNESLSFANI 2322
Qy 87 ----PYLAHLSRMYNGYAGG-VEVQVLLAGNAFTAGK-----ILFAAIP 125
Db 2323 GQTAFYTTTISINIGNTDATNIVFTDVLPLNGVTFVPNTLTVGVLPDANPNTGVLLATLP 2382
Qy 126 PNFVLDMISPAQITMLPHLIVDVRTLEPIMTPLDVRNVVYHFNQF--QPRMLVAMLY 183
Db 2383 PNEIYSIV--FQVT-----VNSIPPI-NPAPNTASTTYEFTVDVNPVPPVLSAATSNT 2431
Qy 184 TPLRSNGSGDDVFTVSCVLTRETTPPEF-----NANIISTKTDLTADVGNTTFTLNLPNTGNVTATDVTVIDILD 2482
Db 2432 TLLQIN-----NANIISTKTDLTADVGNTTFTLNLPNTGNVTATDVTVIDILD 2482
Qy 213 --IYLVPPS-----VESKTPPLPI---LTISELT 238
Db 2483 SNLTFVPNSFTVNGQITLNADLSTGVNIGSINGGTAAIVTFQATVTLPLNNPISNSALT 2542
Qy 239 NSRPEI-PIEQLYTAPNETVNVQNGRCRTLDBELQGTOLLSSAVCFLOCRVAONGD- 296
Db 2543 TYRIYVDPDQSPITTSNGSNTTTQINSAILTAQKSTNVFTVDIGDIVSVTITNSGV 2602
Qy 297 NWDQNLQLTYPNGASYDP-----TDEV 319
Db 2603 NATNVIFTDVIPTDGTSTFEPNSFTLNGTIENANITGVPGDIAPNESAIVEPHITSNEI 2662
Qy 320 PA--PLGTQDFSGMLYGVLTQDNVNVSTGEAKNAGIYISTTSCK---FTPKIGSIGLHS 374
Db 2663 PAINFI-TNQASVSFQHVNPANPPVS---KNITSNSVTTTIESAILTTTKIGDKAFAT 2717
Qy 375 ITEVHPNQSRFTFVGVAVDENT-----PFQOQVLPVHVGSLAENTNLAPAVAPT---- 425
Db 2718 IGDTI--TVTTTITNIGNIPANNVIFSDPIPSWT-QFVAGSVYVDGTPLPASITSGIGI 2774
Qy 426 ----FPGEQL-LFERSRVPQVQLOGQDAFIDCLLPQEWNVHFYQ-----EAPSQADVAL 476
Db 2775 NTIIPNQTVTIIFQ-----QVIVSNPPTFPELQNLAFVNFQYVNGVALQAPQGNVETNV 2829
Qy 477 IRYVNPDTGRTLFEAKLHRSFGITVSHTCAYPLVVPNCHFRFDSWVNOFYSLAPMGT 534
Db 2830 --FVTAHSAILSAVKTAFTASTAFANIGDITTYTVLIONSGNTNATN-VN-FSDLIIFGGT 2883

RESULT 11
US-10-080-334-200
; Sequence 200, Application US/10080334
; Publication No. US20040002584A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shinkets, Richard A
; APPLICANT: Li, Li

APPLICANT: Shenoy, Suresh G
APPLICANT: Kekuda, Ramesh
APPLICANT: Spytek, Kimberly A
APPLICANT: Vernet, Corine A. M.
APPLICANT: Malyankar, Uriel M
APPLICANT: Guo, Xiaojia
APPLICANT: Gusev, Vladimir Y
APPLICANT: Casman, Stacie J
APPLICANT: Boldog, Ferenc L
APPLICANT: Furtak, Katarzyna
APPLICANT: Tchernev, Velizar T
APPLICANT: Patturajan, Meera
APPLICANT: Gangolli, Esha A
APPLICANT: Padigaru, Muralidhara
APPLICANT: Liu, Xiaohong
APPLICANT: Baumgartner, Jason C.
APPLICANT: Gerlach, Valerie
APPLICANT: Spaderna, Steven K
APPLICANT: Zerhusen, Bryan D
TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
FILE REFERENCE: 21402-275
CURRENT APPLICATION NUMBER: US/10/080,334
PRIOR FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: 60/270,523
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/322,712
PRIOR FILING DATE: 2001-09-17
PRIOR APPLICATION NUMBER: 60/311,980
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 60/330,307
PRIOR FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: 60/278,796
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: 60/281,521
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/276,677
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/311,595
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/270,220
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/274,295
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/318,526
PRIOR FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/286,548
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/291,765
PRIOR FILING DATE: 2001-05-17
PRIOR APPLICATION NUMBER: 60/270,797
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/276,400
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/270,810
PRIOR FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 388
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 200
LENGTH: 3317
TYPE: PRT
ORGANISM: Rattus norvegicus
US-10-080-334-200

Query Match 3.9%; Score 111.5; DB 15; Length 3317;
Best Local Similarity 19.2%; Pred. No. 5.1;
Matches 126; Conservative 91; Mismatches 227; Indels 213; Gaps 35;

QY 6 NDAAPSDGAAALVPEGINWPLEPVAG--ASIAAPVAGQTNIIIDPWIRTNFVQAPNGE 63
DB 43 SEDTPGVSGVTQLLARDMDNPLVGVSGEASRFVPEPTGVV--WLRQPLDRETKSE 100
QY 64 FTV--SPRNSPGEIL--LNLELGPDLN-----PYLAHLRMYNGYAGGVEVQVLLA 110

Db 101 FTVEFSVDHQGVITRKVNIQVG-DVNDNAPTFFHNQPSV---RIPENTPVGTPIFVNA 156
QY 111 --GNAFTAGKILFAAIPPN--FLVD-----MISPAQITM----- 140
Db 157 TDPDLGAGSGSVLYSPQSPQFAIDSARGIVTVIRELDYEVYTAQQLTVNATDQDKTRPL 216
QY 141 -----LPHLIVDVRTLEPMTPLPDVRNVFVHNNQPPRMRLVAML-----YT 184
Db 217 STLANLAIIITDVQMDPIFINLPYSTNIYEH--SPPGTTRVITAVDQDKGRPRGIGYT 274
QY 185 PLRNSGSGDDVFTVSC-----RVLTRPTDPE--FIYLV-----PPSVESKTK 225
Db 275 IV--SGNTNSIFALDIYISGALTNLGLDRENPLVSHGHFILLVKGTELNDDRSPSDATVTT 332
QY 226 PFTLPILATIS---ELTNSRFPPIEQI---YTAPNETNVVQCQNGRCTL----- 268
Db 333 TFNILVIDINDNAPEFNSSEYSVAITELAQVGFALPLFIQVVVDKDEGLNSMEFVYLVGN 392
QY 269 -----DGELOGTTQLLSSAVCFIQGRTVADNGNDQNL--QLTYPNGASY----- 313
Db 393 SHHFIISPTSVQKADIRIRVAIPLDYETV---DRYDFDLFANESVDPDHVGYAKVKITL 448
QY 314 -DPTDEVPAPLGTQDFSCMLYGLVLTQDNVNVST-----GEAKNAKGIYIS 357
Db 449 INENDNRPI-----FSQPLYNVSLYENITVGTSLVTLATDNDVGTTEGV-----NYPPS 498
QY 358 TTSGKFTPKIGSIGHSITHEVHPNQQRFTPVGVAVD---ENTPFOQWVLPHYAGSLAL 414
Db 499 DDPDRFSLD-KDTGLIMLIARLDYELIQRFITLTVIARDGSGEET-----TGRVRI 547
QY 415 MTNLAPAVATFPGEQLLFFRSRVPVCGVQLOGQDAFIDCLLPQEWNNHFYQEAAPSQADV 474
Db 548 N-----VLDVNDNVPTFQ---KDAYVGAL-----RENEPSVTQL 578
QY 475 ALIR-----YUNPDTGRTLFEAKLHRS-GFITVSHTGAYPLVVPNG 515
Db 579 VRLRATDESDPPNNLITYSVNASAFGYFDISVYEGYVISVSRPLDYEQI--PNG 633

RESULT 12

US-10-028-248A-74

Sequence 74, Application US/10028248A

Publication No. US20030235882A1

GENERAL INFORMATION:

APPLICANT: Shimkets, Richard

APPLICANT: Patturajan, Meera

APPLICANT: Vernet, Corine

APPLICANT: Casman, Stacie

APPLICANT: Malyankar, Uriel

APPLICANT: Shenoy, Suresh

APPLICANT: Spytek, Kimberly

APPLICANT: Gangolli, Esha

APPLICANT: Miller, Charles

APPLICANT: Boldog, Ferenc

APPLICANT: Li, Li

APPLICANT: Taupier Jr, Raymond J

APPLICANT: Kekuda, Ramesh

APPLICANT: Smithson, Glenda

APPLICANT: Zerhusen, Bryan

APPLICANT: Liu, Xiaohong

APPLICANT: Colman, Steven

APPLICANT: Tchernev, Velizar

APPLICANT: Si, Jingsheng

APPLICANT: Edinger, Shlomit

APPLICANT: Stone, David

APPLICANT: Sciore, Paul

APPLICANT: Millet, Isabelle

APPLICANT: Rothenberg, Mark

TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods

TITLE OF INVENTION: Theroof

FILE REFERENCE: 21402-222

CURRENT APPLICATION NUMBER: US/10/028,248A

```

;
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
; LENGTH: 5376
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-028-248A-74

Query Match 3.9%; Score 110.5; DB 15; Length 5376;
Best Local Similarity 20.9%; Pred. No. 14;
Matches 91; Conservative 56; Mismatches 177; Indels 111; Gaps 20;

QY 26 TMPLP-PVAGASIAAPVAGQNIID--PWIRNFVQAPNGEFTVSPRNS--PGEILLNLE 80
DB 683 TVTPEVPIVLEATAFPTGETTLYTEVPTVPTVTEVTVNVSPEETSVPTETISTE 742
QY 81 LGPDLPYLAHLRMYNGYAGGVEQ-VLLAGNAFTAGKI-LFAAIPP-----NFLV 130
DB 743 V-TTVSPEETVP-----TEVPIVLEATASTGEITLYTEVPTVTEVTVNVSPEETSV 792
QY 131 DMISPAQITMLP---HLIVDVRTLEPIMTLP-----DVRNVFYHNNQOPRMLV 179
DB 793 TNVSPEE-TSVPTETISTEVTTSPEETTLPTTEVPTVSTEVTVNVSPEETSV-PEETIL 850
QY 180 AMLYT-----PLRSGSGDDVFTVSCRVLTPTPDFFETIYLVPPSVESKTKPTLPILT 234
DB 851 TTYLTVPTVPTVTEVTVNVSPEETSVPTETISTEVTTSPEETTLPTTEVPTVSTEV 909
QY 235 SELTNSRFPPIEOLYAPNETNVVQCGNCRCLDGLQGTTLQLLSSAVCFLOQRTVADN 294
DB 910 TEVTN-----VSPETSVPPETILTEI-----TTVSPEETVPPIEGTTL--- 949
QY 295 GDNWDQNLQLTYPNGASYDPTDEVAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKI 354
DB 950 -----PTEVLTVPIEVTTF-----PTGE----- 967
QY 355 YISTSGKFTPKIGS--IGLHSITEHVHPNQSRFTPGVAVDENTPFGQWVLPVHAGSL 412
DB 968 ---TTVPTVPTVSTGMTGVHTVTVTFPEETSIPTVATVLPASIPPEETTTTTEVTTT 1024

;
; RESULT 13
; US-10-07-782-74
; Sequence 74, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steve,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Gangolli, Esha,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Li, Li,
; APPLICANT: Liu, Xiaohong,
; APPLICANT: Malyankar, Uriel,
; APPLICANT: Miller, Charles,
;
;
; APPLICANT: Millet, Isabelle,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Sciore, Paul,
; APPLICANT: Shenoy, Suresh,
; APPLICANT: Shinkets, Richard,
; APPLICANT: Si, Jingsheng,
; APPLICANT: Smithson, Glennda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Stone, David,
; APPLICANT: Taupier, Raymond, Jr.,
; APPLICANT: Tchernev, Velizar,
; APPLICANT: Vernet, Corine,
; APPLICANT: Zerhusen, Brian
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 21402-222CIP
; CURRENT APPLICATION NUMBER: US/10/107,782
; CURRENT FILING DATE: 2002-03-27 248
; PRIOR APPLICATION NUMBER: 10/028,248
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262,959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272,408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285,189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308,039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311,266
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/279,344
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 74
; LENGTH: 5376
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-107-782-74

Query Match 3.9%; Score 110.5; DB 15; Length 5376;
Best Local Similarity 20.9%; Pred. No. 14;
Matches 91; Conservative 56; Mismatches 177; Indels 111; Gaps 20;

QY 26 TMPLP-PVAGASIAAPVAGQNIID--PWIRNFVQAPNGEFTVSPRNS--PGEILLNLE 80
DB 683 TVTPEVPIVLEATAFPTGETTLYTEVPTVPTVTEVTVNVSPEETSVPTETISTE 742
QY 81 LGPDLPYLAHLRMYNGYAGGVEQ-VLLAGNAFTAGKI-LFAAIPP-----NFLV 130
DB 743 V-TTVSPEETVP-----TEVPIVLEATASTGEITLYTEVPTVTEVTVNVSPEETSV 792
QY 131 DMISPAQITMLP---HLIVDVRTLEPIMTLP-----DVRNVFYHNNQOPRMLV 179
DB 793 TNVSPEE-TSVPTETISTEVTTSPEETTLPTTEVPTVSTEVTVNVSPEETSV-PEETIL 850
QY 180 AMLYT-----PLRSGSGDDVFTVSCRVLTPTPDFFETIYLVPPSVESKTKPTLPILT 234
DB 851 TTYLTVPTVPTVTEVTVNVSPEETSVPTETISTEVTTSPEETTLPTTEVPTVSTEV 909
QY 235 SELTNSRFPPIEOLYAPNETNVVQCGNCRCLDGLQGTTLQLLSSAVCFLOQRTVADN 294
DB 910 TEVTN-----VSPETSVPPETILTEI-----TTVSPEETVPPIEGTTL--- 949
QY 295 GDNWDQNLQLTYPNGASYDPTDEVAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKI 354
DB 950 -----PTEVLTVPIEVTTF-----PTGE----- 967
QY 355 YISTSGKFTPKIGS--IGLHSITEHVHPNQSRFTPGVAVDENTPFGQWVLPVHAGSL 412
DB 968 ---TTVPTVPTVSTGMTGVHTVTVTFPEETSIPTVATVLPASIPPEETTTTTEVTTT 1024

;
;
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steve,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Gangolli, Esha,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Li, Li,
; APPLICANT: Liu, Xiaohong,
; APPLICANT: Malyankar, Uriel,
; APPLICANT: Miller, Charles,
;
;
; APPLICANT: Millet, Isabelle,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Sciore, Paul,
; APPLICANT: Shenoy, Suresh,
; APPLICANT: Shinkets, Richard,
; APPLICANT: Si, Jingsheng,
; APPLICANT: Smithson, Glennda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Stone, David,
; APPLICANT: Taupier, Raymond, Jr.,
; APPLICANT: Tchernev, Velizar,
; APPLICANT: Vernet, Corine,
; APPLICANT: Zerhusen, Brian
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 21402-222CIP
; CURRENT APPLICATION NUMBER: US/10/107,782
; CURRENT FILING DATE: 2002-03-27 248
; PRIOR APPLICATION NUMBER: 10/028,248
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262,959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272,408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285,189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308,039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311,266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
; LENGTH: 5376
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-028-248A-74
```


us-09-926-799-9.rapb

Wed Jun 2 09:13:34 2004

Search completed: June 1, 2004, 14:04:44
Job time : 36.648 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:34:56 ; Search time 6.54765 Seconds
(without alignments)
4310.252 Million cell updates/sec

Title: US-09-926-799-9
Perfect score: 2863
Sequence: 1 MKNASNDAPSDNGAASLVP.....VNQFYSLAPMTGNGRRRIQ 542

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1830.5	63.9	539	1 COAT LORDV	P54635 lordedale v
2	1159.5	40.5	546	1 COAT SOUV3	Q04542 southampton
3	340	11.9	2344	1 POLN RHV	P27410 rabbit hemo
4	281.5	9.8	2208	1 POLN MANCV	Q69014 manchester
5	279	9.7	668	1 COAT FCVC6	P27404 feline cali
6	276.5	9.7	671	1 COAT FCVF9	P27406 feline cali
7	271.5	9.5	668	1 COAT FCVF4	P27405 feline cali
8	269.5	9.4	702	1 COAT SMSV1	P36284 san miguel
9	240	8.4	703	1 COAT SMSV4	P36285 san miguel
10	145.5	5.1	2333	1 POLG FMDV1	P03306 f genome po
11	130	4.5	2332	1 POLG FMDVA	P03308 f genome po
12	129	4.5	2193	1 POLG CX16T	Q9qf31 c genome po
13	123.5	4.3	746	1 RHTA RHIME	Q923q5 rhizobium m
14	118	4.1	2194	1 POLG HE701	P2537 h genome po
15	116.5	4.1	2193	1 POLG CX16G	Q65900 c genome po
16	116	4.1	545	1 GVPD HALME	Q02229 halobacteri
17	114.5	4.0	725	1 SP3 MOUSE	O70494 mus musculus
18	114	4.0	2336	1 POLG FMDVZ	P49303 f genome po
19	111.5	3.9	3317	1 CADN RAT	P58365 ratus norv
20	111	3.9	1034	1 ADD DROME	P4362 drosophila
21	110.5	3.9	5376	1 ZAN MOUSE	O88799 mus musculus
22	110	3.8	1196	1 XPG XENLA	P14629 xenopus lae
23	110	3.8	2290	1 POLG EMCV	P03304 e genome po
24	109.5	3.8	808	1 POLG HPVAV	Q02381 hepatitis a
25	109.5	3.8	852	1 POLG HPVAVC	P06442 hepatitis a
26	109.5	3.8	2193	1 POLG HE71M	Q66479 h genome po
27	108.5	3.8	901	1 POLG ENMG3	P32540 mengo encep
28	108.5	3.8	2226	1 POLG HPVAV2	P26580 hepatitis a
29	108.5	3.8	2226	1 POLG HPVAV	P26581 hepatitis a
30	108.5	3.8	2227	1 POLG HPVAVH	P08617 hepatitis a
31	108.5	3.8	2227	1 POLG HPVAVL	P06441 hepatitis a
32	108.5	3.8	2227	1 POLG HPVAVM	P13901 hepatitis a
33	108.5	3.8	2230	1 POLG HPVAVS	P14553 simian hepa

```

RESULT 1
COAT LORDV STANDARD; PRT; 539 AA.
AC P54635;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Coat protein (Capped protein).
OS Lordedale virus (Human enteric calicivirus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=82658;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96005060; PubMed=7561776;
RA Dingle K.E., Lambden P.R., Caul E.O., Clarke I.N.;
RT "Human enteric Caliciviridae: the complete genome sequence and
RT expression of virus-like particles from a genetic group II small
RT round structured virus".
RL GenBank:U01766;234952355439594
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALCIVIRUSES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; X86557; CAA60255.1;
CC InterPro; IPR004005; Calici_coat.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00915; Calici_coat; 1.
CC
CC Coat protein; Glycoprotein.
CC CARBOHYD 172 472 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 531 531 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 539 AA; 58775 MW; 5E5C63E7F2C5FD21 CRC64;
Query Match 63.9%; Score 1830.5; DB 1; Length 539;
Best Local Similarity 64.0%; Pred. No. 4e-129;
Matches 349; Conservative 73; Mismatches 110; Indels 13; Gaps 6;
QY 1 MKNASNDAPSDNGAASLVPESGINTPLEPVAGASIAAPVAGQTNIDPWIRTFVQAP 60
Db 1 MKNASNDAPSDGSAANLVPEVNEVMALEPVVGAIAAPVAGQNVIDPWIRNFVQAP 60
QY 61 NGEFTVSPRNSGPEILLNLELGPDLNPLVLAHLSRMVYAGVGVQVLLAGNAFTAGKTL 120
Db 61 GGEFTVSPRNAPGELWSAPLGPDLNPLVLSHLSRMVYAGVGVQVLLAGNAFTAGKVI 120
QY 121 FAATIPPNFLVDMISPAQITMLPHLIVDVTLEPINTPLPDVNVFVHNNQPPRMRLVA 180
Db 121 FAAPVPNPFTEGLSPQVTMFPHIIVDRQLEFVLIPLDVVRNRFVHYNQANDSTKLKLA 180

```

ALIGNMENTS

```

RESULT 1
COAT LORDV STANDARD; PRT; 539 AA.
AC P54635;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Coat protein (Capped protein).
OS Lordedale virus (Human enteric calicivirus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=82658;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96005060; PubMed=7561776;
RA Dingle K.E., Lambden P.R., Caul E.O., Clarke I.N.;
RT "Human enteric Caliciviridae: the complete genome sequence and
RT expression of virus-like particles from a genetic group II small
RT round structured virus".
RL GenBank:U01766;234952355439594
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALCIVIRUSES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; X86557; CAA60255.1;
CC InterPro; IPR004005; Calici_coat.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00915; Calici_coat; 1.
CC
CC Coat protein; Glycoprotein.
CC CARBOHYD 172 472 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 531 531 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 539 AA; 58775 MW; 5E5C63E7F2C5FD21 CRC64;
Query Match 63.9%; Score 1830.5; DB 1; Length 539;
Best Local Similarity 64.0%; Pred. No. 4e-129;
Matches 349; Conservative 73; Mismatches 110; Indels 13; Gaps 6;
QY 1 MKNASNDAPSDNGAASLVPESGINTPLEPVAGASIAAPVAGQTNIDPWIRTFVQAP 60
Db 1 MKNASNDAPSDGSAANLVPEVNEVMALEPVVGAIAAPVAGQNVIDPWIRNFVQAP 60
QY 61 NGEFTVSPRNSGPEILLNLELGPDLNPLVLAHLSRMVYAGVGVQVLLAGNAFTAGKTL 120
Db 61 GGEFTVSPRNAPGELWSAPLGPDLNPLVLSHLSRMVYAGVGVQVLLAGNAFTAGKVI 120
QY 121 FAATIPPNFLVDMISPAQITMLPHLIVDVTLEPINTPLPDVNVFVHNNQPPRMRLVA 180
Db 121 FAAPVPNPFTEGLSPQVTMFPHIIVDRQLEFVLIPLDVVRNRFVHYNQANDSTKLKLA 180

```

Matches 258; Conservative 84; Mismatches 176; Indels 51; Gaps 17;

QY 1 MLYTPIRNSGDDVFTVSCVRLTRPTDFFIYLVPPSVESKTKPFTLPILITSELTNS 240
 DB 1 MLYTPIRNSGDDVFTVSCVRLTRPTDFFIYLVPPSVESKTKPFTLPILITSELTNS 240
 QY 241 RPIPIEQLTYTAPNTNVCQNGRCTLGDELQGTLLSSAVCPQLQRTVADNGDNWDQ 300
 DB 241 RPIPIEQLTYTAPNTNVCQNGRCTLGDELQGTLLSSAVCPQLQRTVADNGDNWDQ 300
 QY 301 NLLQTYTPNGASYDTPDPAQLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGIYISTTS 360
 DB 299 YTMNLASQWNSYDTEEPAPLGTDPFVGKIQGLLTQTRADGSTRAHKAT---VSTGS 355
 QY 361 GKFTPKIGSIGLHSITEH-VHNPQOSRFTFPVGAVD-----ENTPQQWVLPHYAGSLAIN 415
 DB 356 VHTFPKLSGVQFTDNNDFQAGQNTKFTFPVGIQDGDHQNKP-QQWSLPNYSGRTHN 414
 QY 416 TNLAPAVATPFGEOQLLFRSRVPCVOGQDQDAFDCLLPQEWVNFHFOEAPQSDVA 475
 DB 415 VHLAPAVATPFGEOQLLFRSRVPCVOGQDQDAFDCLLPQEWVNFHFOEAPQSDVA 472
 QY 476 LIRYVNPDTGRTLFEAKLHRSGLTIVSHTGAYPLVVPNGHFRFDSWVNFQFYSLAPMGTG 535
 DB 473 LLRFVNPDTGRTLFEAKLHRSGLTIVSHTGAYPLVVPNGHFRFDSWVNFQFYSLAPMGTG 532
 QY 536 NGRRR 540
 DB 533 TGRRR 537

RESULT 2
 COAT_SOUV3 STANDARD; PRT; 546 AA.
 ID COAT_SOUV3
 AC Q04542:
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Coat protein (Capsid protein).
 OS Southampton virus (serotype 3).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 OC NCBI_TaxID=37129;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93142023; PubMed=8380940;
 RA Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;
 RT "Sequence and genome organization of a human small round-structured
 RT (Norwalk-like) virus."
 RL Science 259:516-519(1993).
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; L07418; AAA92984.1; -.
 DR PIR; B37491; B37491.
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 546 AA; 58774 MW; C82B2A85AD4B05EA CRC64;
 Query Match 40.5%; Score 1159.5; DB 1; Length 546;
 Best Local Similarity 45.3%; Pred. No. 5.2e-79;

MATCHES 258; CONSERVATIVE 84; MISMATCHES 176; INDELS 51; GAPS 17;

QY 1 MLYTPIRNSGDDVFTVSCVRLTRPTDFFIYLVPPSVESKTKPFTLPILITSELTNS 240
 DB 1 MLYTPIRNSGDDVFTVSCVRLTRPTDFFIYLVPPSVESKTKPFTLPILITSELTNS 240
 QY 241 RPIPIEQLTYTAPNTNVCQNGRCTLGDELQGTLLSSAVCPQLQRTVADNGDNWDQ 300
 DB 241 RPIPIEQLTYTAPNTNVCQNGRCTLGDELQGTLLSSAVCPQLQRTVADNGDNWDQ 300
 QY 301 NLLQTYTPNGASYDTPDPAQLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGIYISTTS 360
 DB 299 YTMNLASQWNSYDTEEPAPLGTDPFVGKIQGLLTQTRADGSTRAHKAT---VSTGS 355
 QY 361 GKFTPKIGSIGLHSITEH-VHNPQOSRFTFPVGAVD-----ENTPQQWVLPHYAGSLAIN 415
 DB 356 VHTFPKLSGVQFTDNNDFQAGQNTKFTFPVGIQDGDHQNKP-QQWSLPNYSGRTHN 414
 QY 416 TNLAPAVATPFGEOQLLFRSRVPCVOGQDQDAFDCLLPQEWVNFHFOEAPQSDVA 475
 DB 415 VHLAPAVATPFGEOQLLFRSRVPCVOGQDQDAFDCLLPQEWVNFHFOEAPQSDVA 472
 QY 476 LIRYVNPDTGRTLFEAKLHRSGLTIVSHTGAYPLVVPNGHFRFDSWVNFQFYSLAPMGTG 535
 DB 473 LLRFVNPDTGRTLFEAKLHRSGLTIVSHTGAYPLVVPNGHFRFDSWVNFQFYSLAPMGTG 532
 QY 536 NGRRR 540
 DB 533 TGRRR 537

RESULT 3
 POLN_RHDV STANDARD; PRT; 2344 AA.
 ID POLN_RHDV
 AC P27410;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Non-structural polyprotein (Contains: RNA-directed RNA polymerase
 DE (EC 2.7.7.48); Thiol protease P3C (EC 3.4.22.-); Helicase (2C like
 DE protein); Coat protein).
 OS Rabbit hemorrhagic disease virus (RHDV).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Lagovirus.
 OC NCBI_TaxID=11976;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91361557; PubMed=1840711;
 RA Meyers G., Wirblich C., Thiel H.-J.;
 RT "Rabbit hemorrhagic disease virus -- molecular cloning and nucleotide
 RT sequencing of a calicivirus genome."
 RL Virology 184:684-676(1991).
 CC -!- FUNCTION: THE P3C CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE
 CC FOR CLEAVAGE AT CERTAIN Q/G SITES IN THE POLYPROTEIN.
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA) (N).
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M67473; AAA47285.1; -
DR PIR; A41039; BRWEH
DR PDB; 1KHV; 16-JAN-02.
DR PDB; 1KHV; 16-JAN-02.
DR MEROPS; C24.001; -
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR004005; Calici coat.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR000317; Peptidase_C24.
DR InterPro; IPR000605; RNA helicase.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR001205; RNA pol p3d.
DR InterPro; IPR007094; RNA_pol_Psvir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR Pfam; PF03510; Peptidase_C24; 1.
DR Pfam; PF00680; RNA dep RNA_pol; 1.
DR Pfam; PF00910; RNA helicase; 1.
DR PRINTS; PR00916; 2CENDOPTASE.
DR PRINTS; PR00918; CALICIVIRUSNS.
DR SMART; SM00382; AAA; 1.
KW Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;
KW Thiol protease; Helicase; ATP-binding; Coat protein; 3D-structure.
FT CHAIN 7508 7707
FT CHAIN 1109 1251
FT CHAIN 71497 71625
FT CHAIN ? 2344
FT CHAIN ? 2344
FT NP_BIND 522 539
FT ACT_SITE 1212 1212
FT ACT_SITE 1227 1227
FT ACT_SITE 1227 1227
SQ SEQUENCE 2344 AA; 257066 MW; 1454C248F81E9212 CRC64;

Query Match 11.9%; Score 340; DB 1; Length 2344;

Best Local Similarity 27.5%; Pred. No. 5.9e-17;

Matches 145; Conservative 72; Mismatches 213; Indels 98; Gaps 25;

QY 8 AAPSNDGAASL-----VPEGINTEPLEPVA-----GASIA-APVAGQTNID--- 49
DB 1772 AAQGEAAAGTATTASVPGTTDGMDPGVVATTSVITAENSSAIATAGTGGPQQVDQGE 1831
QY 50 PWIRTNFOAPNGEFTVSPRNSPGEILLNLEGLDPLNPLYLAHLSRMVYAGVEVQVLL 109
DB 1832 TW-RNFYY--NDVFTWSVADAPGSLITYVQHSPPQNPPTAVLSQNYAGWAGMGRFIV 1888
QY 110 AGNAFTAGKILFAALPPNPLVDIMISPAQITMLPLHIVDVTLEIMTLPDVRNVFYHFN 169
DB 1889 AGSGVFGGLVRVAVIPPG--IEIGPGLVVRQFPVHVVIDARSLPEVITITMPDLRPNMYHT 1946
QY 170 NOPQRMRLVAMLYPLRNSGDDVFTVSCVLRTPDPDFEFLVLPVSVEKTKPFTL 229
DB 1947 GDPLVTLVLVSYYNL-INPFGSGTSATQVTVTRPSDDFEFVMIRAPS--SKTVDSIS 2003
QY 230 P--ILTISLT---NSRPIPIEOLYTAPNETNVVQCONGRCTLDGELQGTTLSSAV 283
DB 2004 PAGLUTTPVLTGVGNDNRWNGQIVGLQVPVGGFST--C-NRHNUNGSTYGWSSPRFGDI 2060
QY 284 CFLQRTVADNGDNDQNLQLTYPNGASY--DPTDEVPAPLGTQD-----FSG----- 330
DB 2061 DHRRG-SASYSGSN-ATNVLPQWYANAGSAIDNPISQV-APDGFDPMSVPVNGGPIAA 2117
QY 331 --MLGVLTQDN---VNVSTGEAKN-----AKGIYISTT--- 359
DB 2118 GWVGFCAIWNNSGAPNVTTVQAYELGATGAPGNLQPTTNTSGATVAKSIYAVVTGTA 2177
QY 360 ---SGKFTPKIGSLGHSITEVHVNQOQSRF-----TPGVAVDENTPQQWVLPYAGS 411
DB 2178 QNPAGLVFNWASGIISTPNASAITYTPQDPRIVTTPTGTPAAAPVGKNTPIM-----PAS 2230

QY 412 LALNTNLAPAVAPTTPGGQLLFFRSRVPVCVQGLQGQDAPIDCLLPQEW 459
DB 2231 VVRRGTGVDNATAGSANGTYGTGSQLPVTIGL-SLNNYSSALMPEGQF 2277

RESULT 4

POLN_MANCV STANDARD; PRT; 2208 AA.
ID POLN_MANCV AC Q69014;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: RNA-directed RNA polymerase
DE (EC 2.7.7.48); Thiol protease 3C (EC 3.4.22.-); Helicase (2C like
DE protein); Coat protein].
OS Manchestera virus (Human enteric calicivirus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Sapovirus.
OX NCBI_TaxID=82659;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95390791; PubMed=7661689;
RA Liu B.L., Clarke I.N., Caul E.O., Lambden P.R.;
RT "Human enteric caliciviruses have a unique genome structure and are
RT distinct from the Norwalk-like viruses."
RL Arch. Virol. 140:1345-1356(1995).
CC -!- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).
CC -!- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR
CC THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X86560; CAA60262.1; -
DR MEROPS; C24.UPW; -
DR InterPro; IPR004005; Calici coat.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR000317; Peptidase_C24.
DR InterPro; IPR000605; RNA helicase.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR001205; RNA pol p3d.
DR InterPro; IPR007094; RNA_pol_Psvir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR Pfam; PF03510; Peptidase_C24; 1.
DR Pfam; PF00680; RNA dep RNA_pol; 1.
DR Pfam; PF00910; RNA helicase; 1.
DR PRINTS; PR00916; 2CENDOPTASE.
DR PRINTS; PR00918; CALICIVIRUSNS.
KW Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;
KW Thiol protease; Helicase; ATP-binding; Coat protein.
FT CHAIN ? ?
FT CHAIN ? ?
FT CHAIN ? ?
FT CHAIN ? ?
FT NP_BIND 408 415
FT ACT_SITE 1097 1097
FT ACT_SITE 1112 1112
FT ACT_SITE 1112 1112
SQ SEQUENCE 2208 AA; 242736 MW; 3E299D5BA20E45DC CRC64;

Query Match 9.8%; Score 281.5; DB 1; Length 2208;

Best Local Similarity 21.7%; Pred. No. 1.3e-12;

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

Matches 128; Conservative 92; Mismatches 228; Indels 143; Gaps 22;	
QY	20 PEGINETMPLFVAGASIAAPVAGQTNIIIDPWIRTNFVQAPNGEFTVSPRNSGEILLNL 79
Db	1687 PNGAQRLELAATGAI-----QSNVPEA-IRNCF--AVFRTFAWDRMPTGTFGLGSI 1736
QY	80 ELGPDNLNPLYLAHLSRWYNGYAGGVQVLLAGNATAGKILFAAIPNPNFLVDMISPAQIT 139
Db	1737 SLHPNINPYTSLSGMAGWGSGFEVRLSGSGVFAGRITIASVIPPG--VDPSSIRDPG 1794
QY	140 MLPHLIVDVRTLEPIMTLPVDRNRYFHNQOPQPRMELVAMLYTPLNSGSGDDVTVTS 199
Db	1795 VLPHAFVDAIRITPEVSMIPDVAVDYHRMDGAETCSLGEVWVQPLNPFSTTAVSTCW 1854
QY	200 CRVLTRPTDFEFTIYVLP--SVESKTKP-----FTLPIITISE--L 237
Db	1855 VSVETKPGGDFDCLLRPFQGMENGVSPEGLLPRRLGYSRGNRVGGLVGMILVAEHKQ 1914
QY	238 TNSRF-----PIEQLYTAPNETN-----VVOQN-----G 264
Db	1915 VNRHNSNVTFCWSTAPVNPMAAEIVTQAHSTSRHAWLSIGAQNKGPLFFGIPNHPFD 1974
QY	265 RC-----TLDGELQ--TTOLLSSAVCFLOQRTVADGNDQNLLOLTYPNGASYDPT 316
Db	1975 SCATVVGMADTSLGGRPSTGVCGPAISF-----QNGGDVYENDTPPSVMF--ATYDPL 2025
QY	317 DEVPAPIGTQDFSGMLYGLVLTQDNVNVSTGEAKNAKGIYISTTSKGFTPKIGSLGHSIT 376
Db	2026 TSGTGVALTNSINPASILVRIENDDPTSGFANDKXNVVVQMSWEMVT-----GTNQIR 2079
QY	377 EHVHPNQSRFTPVGVAVDENTFPQWVLPHYAGSLALNTNLAPAVAPTFPGEOALLFES 436
Db	2080 GQVTPMSGTNYTFTSTGANTLVLMQERML-----SYDGHQOALLYS 2120
QY	437 RVPVQGLQGDQAFIDCLLQEWVNYHFYQEAAPSQADVALIRYVNPDTGRTLFEAKLHRS 496
Db	2121 QL-----ERTAEYFQNDIVNIPNSMAVF--NVEITNSASFQIGIRPD 2160
QY	497 GFITVSHTGAYPLVVPNGHFRFDSWVNYQFVSLAPMG-----TGN-GRRR 540
Db	2161 GYVVTG--GSIGVNVLEPETRF-----QYVGLPLSALSGPSGNGRAR 2204
RESULT 5	
COAT_FCV66	
ID	COAT_FCV66 STANDARD; PRT; 668 AA.
AC	P27404;
DT	01-AUG-1992 (Rel. 23, Created)
DT	01-AUG-1992 (Rel. 23, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Coat protein (Capsid protein).
OS	Feline calicivirus (strain CFI/68 FIV) (FCV).
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae; Vesivirus.
OX	NCBI_TaxID=11979;
RN	[1]
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX	MEDLINE=91374597; PubMed=1716692;
RA	Neill J.D., Reardon I.M., Heinrikson R.L.;
RT	"Nucleotide sequence and expression of the capsid protein gene of feline calicivirus."
RL	J. Virol. 65:5440-5447(1991).
RN	[2]
RP	SEQUENCE FROM N.A.
RA	Neill J.D.;
RL	Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC	-1- PTM: The N-terminus is blocked.
CC	-1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

Query Match 9.7%; Score 279; DB 1; Length 668;	
Best Local Similarity 23.2%; Pred. No. 3.8e-13;	
Matches 137; Conservative 81; Mismatches 210; Indels 162; Gaps 28;	
QY	11 SNGGAASLVPEGINETMP-----LEPVAGASIAAPVAGQTNIIIDPW-IRTNFVQAPNGEFT 65
Db	125 ADDGSITTEPQG---TMVGGVIAEPNAQMSAADWATGKSVDSWEAFSPHSTVNW--- 178
QY	66 VSPRNSPGEILLNLEGLPDNLNPLYLAHLSRWYNGYAGGVQVLLAGNATAGKILFAAIP 125
Db	179 -STSETQKILFKQSLGPLLNPYLTHAKLYVAVSGSVDRFSISGSGVFGKJAAIVVP 237
QY	126 PNFVDMISPAQITMLPHLIVDVRTLEPIMTLPVDRNRYFHNQOPQPRMELVAMLYTP 185
Db	238 PG--IDPVQSTMSLQVPHVLFDAQVPEPVIFSPIDLRSLYHLMSD--TTSIVIMVYND 294
QY	186 L-----RNSGSGDDVFTVSCRYLTPRTPDFEFTIYVLP-----SVESK----- 223
Db	295 LINPYANDSNSGCV-----TVETKPGDPFKHLLKPPGSMTHGSIPLSDLIPIKSSSLW 349
QY	224 -----TKPFTL-----PI-LTISELNSRFPPIEQ 248
Db	350 IGNRFWSDITDFVIRPFPVQANRHDFNQTAGMSTFRPRPITITISVKESAKLGIGVAT 409
QY	249 LYTAPN-----ETNVVQCNGRCCTLDGELQGTITOLLSSAVCFLOQRTVADGNDQNL 302
Db	410 DYIVEGIPDGMPDT-----TIPGELVPV-----GDYAITNGTND-- 444
QY	303 LQLYTPNGASYDPTDEVPAPLGTQDFSGM-LYGVLTQ-----DNVNVSTGEAKNAKG 353
Db	445 -----ITTAQYDAATEI---RNNTNFRMYTCGSLQRAWGDKKISNTAFITGIVDQAKL 497
QY	354 IYISTSGKFTPKIGSIGLHSTIEHVHNQOS-----RFTPVG---VAVDENTFPQWVLP 406
Db	498 IPSNTID---QTKIAVFOOTHANKHVQTSDDTLALLGYTGIGEEAIGADRVRVIRISVLP 554
QY	407 HYAGSLALNTNLAPAVAPTFPGEOQLLFRSVPVCFQGLQGDQAFIDCLL-----PQEWVNH 463
Db	555 ERGAR-----GGNHDFHFKNSIKLGYVIRSDVFNSTQLHSTSRQLSLNHY 599
QY	464 YQBAAPSQADVALIRYVNPDTGRTLFEAKLHRS--FTVSHTG--AYPL 509
Db	600 LL-----SPDSFAVYRII--DSNGSWPFDIGINDGFSFGVSGSIGKLEFPL 643
RESULT 6	
COAT_FCV69	
ID	COAT_FCV69 STANDARD; PRT; 671 AA.
AC	P27406;
DT	01-AUG-1992 (Rel. 23, Created)
DT	01-AUG-1992 (Rel. 23, Last sequence update)
DT	01-JUN-1994 (Rel. 29, Last annotation update)
DE	Coat protein (Capsid protein).
GN	CCP76.

OS Feline calicivirus (strain F9) (FCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OX NCBI_TaxID=11981;
 RN [1]
 RP SEQUENCE FROM N.A. MEDLINE=92410623; PubMed=1529544;
 RX Carter M.J., Milcon I.D., Meanger J., Bennett M., Gaskell R.M.,
 RA Turner P.C.;
 RT "The complete nucleotide sequence of a feline calicivirus";
 RL Virology 190:443-448(1992).
 RN [2]
 RP SEQUENCE FROM N.A. MEDLINE=1633955;
 RX Meanger J., Carter M.J., Gaskell R.M., Turner P.C.;
 RA "Cloning and sequence determination of the feline calicivirus strain
 F9";
 RL Biochem. Soc. Trans. 20:265-268(1992).
 CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; M86379; AAA79327.1; -;
 DR EMBL; Z11536; CA77636.1; -;
 DR PIR; B43382; VCMWF9.
 DR InterPro; IPR004005; Calici_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 618 618 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 671 AA; 73441 MW; 338E86D8370D5E5 CRC64;
 Query Match 9.7%; Score 276.5; DB 1; Length 671;
 Best Local Similarity 25.1%; Pred. No. 5.8e-13;
 Matches 131; Conservative 68; Mismatches 171; Indels 151; Gaps 28;
 QY 11 SNDGAASLVPEGINETMP---LEPVAGASTAAPVAGQTNIDPW-IRTNFVQAPNGEFT 65
 DB 125 ADDGSITAEQG---TWVGGVIAEPSAQMSTAAADMATKGSVDSEWFAFFSHTSVNW--- 178
 QY 66 VSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGGVQVLLAGNAFTAGKILFAAIP 125
 DB 179 -STSETQKILFKQSLGPLLPYLAHLSRMVYAGGVQVLLAGNAFTAGKILFAAIP 237
 QY 126 ENFLVDMISPAQITMLPHLIVDRTLEPIMTLPDVRNVFHNQOPQPMRLVAMLYTP 185
 DB 238 PG--VDPVQSTSMQLQYPHVLFDAQVEPVIFCLPDLRSTLYHLSMDTD--TTSLVIMVYND 294
 QY 186 LRSNGSGDDVFTVSC--RVLRTPDPEFIYVLP-----SVES-----KT----- 224
 DB 295 L-INPYANDANSSGCVITVETKPGDPFKHLKPPGSMVLTGHSIPSLIPKTSLSLWIGNR 353
 QY 225 -----KPFLL-----PI-LTISELTSNRFPPIPIQLYTA 252
 DB 354 YWSDITDFVIRPFVQANRHFDFNQETAGWSTPRPRPISVITEQNGAKLIGVATDIV 413
 QY 253 PN-----ETNVQCONGRCTLDGEL-----QGTQLLSSAVCFQGRVAD---N 294
 DB 414 PGIPDGPDT-----TIPGELIPAGDYAITNGTGNDITTAT---GYDTADIKN 459
 QY 295 GDN-----W-DQMLLQITYNGASYD--PTDEVPAPLGTQFSGMLYGLVLTQD 339
 DB 460 NTNFRMGYICGSLQRAWGDKISNTAFITTATLDGDNNNKINPCNTIDQSKI---VFPQD 516

QY 340 N-----VNVG-----TGEAKNAKG-----IYISTSGKFTPKIGSLGHSITHV 379
 DB 517 NHVGKKAQTSDDTLALLGYTGIGGQAIGSDRDRVVRIST-----LPETGARGGN----- 565
 QY 380 HPNQOSRFTPGVAVDENTPFQQWVLPVHAGSLAINTNLAP 420
 DB 566 HPIFYKNSIKLGVYIRSIDVFNQSL-HTSQSLNHYLLP 605

RESULT 7

COAT_FCVF4
 ID_COAT_FCVF4 STANDARD; PRT; 668 AA.
 AC F27405;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Coat protein (Capsid protein).
 OS Feline calicivirus (strain Japanese F4) (FCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OX NCBI_TaxID=11980;
 RN [1]
 RP SEQUENCE FROM N.A. MEDLINE=1853578;
 RX Miyamura K., Taniguchi Y., Takahashi E., Utagawa E., Takeda N.,
 RA Miyamura K., Yamazaki S., Mikami T.;
 RT "Sequence analysis of the 3'-end of feline calicivirus genome";
 RL Virology 183:810-814(1991).
 CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; D90357; BAA14371.1; -;
 DR PIR; B40481; VCMWF9.
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 668 AA; 73589 MW; 85BBD8C85804E503 CRC64;

Query Match 9.5%; Score 271.5; DB 1; Length 668;

Best Local Similarity 22.5%; Pred. No. 1.4e-12;
Matches 130; Conservative 77; Mismatches 193; Indels 179; Gaps 25;

QY 30 EPVAGASTAAPVAGQTNIDPW-IRTNFVQAPNGEFTVSPRNSPGEILLNLELGPDLNPLY 88
 DB 145 EPSAQMSTAAADMATKGSVDSEWFAFFSHTSVNW-----STSETQKILFKQSLGPLLPY 200
 QY 89 LAHLSRMVYAGGVQVLLAGNAFTAGKILFAAIPENFLVDMISPAQITMLPHLIVDV 148
 DB 201 LEHLSKLYVANSVGSIEVRFSISGSGVFGKLAIVVPPG--VDPVQSTSMQLQYPHVLFDA 258
 QY 149 RTLEPIMTLPDVRNVFHNQOPQPMRLVAMLYTP-----RNSGSGDDVFTVSCR 202
 DB 259 RQVEFVITPDLRSTLYHLSMDTD--TTSLVIMVYNDLINPYANDSNSSGCV-----TV 312
 QY 203 LTRPTDPEFIYVLP-----SVES-----TKPFTL----- 229
 DB 313 ETKPGDPFKHLKPPGSMVLTGHSIPSLIPKSSLSLWIGNRYWTDITDFVIRPFVQANR 372

QY 230 -----PI-LTISELNSRPPIEOLYTPAPNETNVVQCGRCITLDGE 271
 DB 373 HFDENQETAGWSTPRFRPITITISEKNGKLGIGVADYIIPGP----- 417
 QY 272 LOGITQLLSSAVCFLOGRTVADNGDNW-DONLLQLYTPNGASYDPTDEVPAPIGTQDFSG 330
 DB 418 -----DGMPTTIAKLLIPAGD----- 434
 QY 331 MLYGVLTQDNVNVSTGEAKNAGIYISTTSKGFPKIGSI-----GLHSIT----- 376
 DB 435 --YSITTGEGNDIKTAQAYDAVAAVKNTNFRGMVIGSLORAWGDKKISNTAFITAIR 492
 QY 377 --EHVHPNQSRFTPVGVAVDENTPFOQWV-----LPHYA--GSLALNTNLAPAV-- 422
 DB 493 DGNELKPSWTIMTK--LAVYODTHVEQVQTSDDTLALLGYTGIGEAIGNSRDRVRI 550
 QY 423 -----APTFFGQLLFFRSRVPVCOGLOQDAFIDCLL---POENVHFFYQEAAPSQADV 474
 DB 551 SVLPEAGARGGNHPIFYKNSIKLGIVIRSIDVFNFSQILHTRQRLSLNHY---LLPPDS-F 606
 QY 475 ALIRYVNPDTGRTLPEAKLHRSQ--FITVSHTG--AYPL 509
 DB 607 AVYRII--DSNGSWFDIGIDSGFSEFVGVSIDIGKLEFPL 643
 RESULT 8
 COAT_SMSV1 STANDARD; PRT; 702 AA.
 AC P36284;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Coat protein (Capsid protein).
 OS San Miguel sea lion virus (serotype 1) (SMSV 1).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OX NCBI_TaxID=36406;
 RN SEQUENCE FROM N.A.
 RP MEDLINE=92410750; PubMed=1529644;
 RX Neill J.D.;
 RA "Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lion virus: identification of conserved and non-conserved amino acid sequences among calicivirus capsid proteins.";
 RT Virus Res. 24:211-222(1992).
 RL -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; M87481; AAA16217.1; -.
 DR PIR; A48562; A48562. Calici coat.
 DR InterPro; IPR004005; Calici coat.
 DR Pfam; PF00915; Viral_Cap_coat.
 DR Coat protein; Glycoprotein.
 KW CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 493 493 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SEQUENCE 702 AA; E6E5A58523DEE3D7 CRC64;
 SQ
 Query Match 9.4%; Score 269.5; DB 1; Length 702;
 Best Local Similarity 25.0%; Pred. No. 2.1e-12;
 Matches 108; Conservative 58; Mismatches 169; Indels 97; Gaps 21;
 QY 8 AAFSNDGAASLVPEGINETMPELVAGAS-----IAAPVAGQNIID-PWIRTFVQ---A 59

DB 151 AESDPPGADIIVTEEQGTWVQQQPVPAQSAALTTAAASTGKT--VDCW--TTFFSYHTA 206
 QY 60 PNGEFTVSPRNSPCEILLNLLELGPDLNPYLAHLRSMYNGVAGGVEQVLLAGNAFTAGKI 119
 DB 207 VNMSTT-----BAQKILFSRALSPELNPYLRHSSLYSTWSGGIDVRFVTSVSGVFGGKL 262
 QY 120 LFAALPPNPLFVDMISPAQITWMLPHLIVDVRTLEIMPTLPDVRNVFHFNNQPPRNLV 179
 DB 263 AALIVPPG--TEPVESPTMLQYPHVLFDARQTEVFIPTDIRKTLX-SMDDTDTTTLV 319
 QY 180 AMLYTPL-RNSGSGDDVTVSCRVLTRTPDPEFYLVP-----SVESKTKP----- 226
 DB 320 INVYNELINPYEQSEPKSSCSITVETRPSSDTFSLKPKGSLKKGISPSDLIPNSRH 379
 QY 227 -----FTLPLTIS-----ELNSRFPPIEOLYTPAPNETNVVQCGRCIT 267
 DB 380 WNGNRWSTIDGFVQVPRVFSQNRHFDSDTTGWSPT-----YVPIEVLTKLDRG--- 432
 QY 268 LDGELGGTTQLLSSAVCFLOGRTVADNGDNWONLLQLLYPNGASYDPTDEVPAPIGTQD 327
 DB 433 --GOYFKVTDTEKSLVPGLP-----DGMPTT-----PT-AMTASNGYD 470
 QY 328 FSCMLYGVLTQDNVNVSTGEAKNAGIYISTTSKGFPKI-GSIGLHSHITHVHNNQSR 386
 DB 471 YTVAYRI-----TNNGTHFKGYI--MGNLTTKVKGSDNLGETQQ---TSRTL 514
 QY 387 FTPVGVAVDENT 398
 DB 515 FASVGNVYKQNT 526
 RESULT 9
 COAT_SMSV4 STANDARD; PRT; 703 AA.
 AC P36285;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Coat protein (Capsid protein).
 OS San Miguel sea lion virus (serotype 4) (SMSV 4).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OX NCBI_TaxID=36407;
 RN SEQUENCE FROM N.A.
 RP MEDLINE=92410750; PubMed=1529644;
 RX Neill J.D.;
 RA "Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lion virus: identification of conserved and non-conserved amino acid sequences among calicivirus capsid proteins.";
 RT Virus Res. 24:211-222(1992).
 RL -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; M87482; AAA16220.1; -.
 DR PIR; C48562; C48562. Calici coat.
 DR InterPro; IPR004005; Calici coat.
 DR InterPro; IPR008975; Viral_Cap_coat.
 DR Pfam; PF00915; Calici coat; 1.
 DR Coat protein; Glycoprotein.
 KW CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SEQUENCE 703 AA; 77721 MW; C5DAD8223B261073 CRC64;
 SQ

Query Match 8.4%; Score 240; DB 1; Length 703;
 Best Local Similarity 29.3%; Pred. No. 3.3e-10;
 Matches 63; Conservative 45; Mismatches 91; Indels 16; Gaps 8;

QY 8 AAPSNDGAASLVEG-----INSTMPLPVPAGASIAAPVAGQNIIDPWIRTNFVQAPNCE 63
 DB 151 AESDQGSABIVTEBQTVVQQPAPAPATALATASTGK-SVQEWMTFF--SYHTS 206
 QY 64 FTVSPRNSPEILLNLELGPDLNPNYLAHLSRMNGYAGGVEVQVLLAGNAFTAGKILFAA 123
 DB 207 INWSTVESQKILYSQALNPSINPYLDHIAKLYSTWSGGIDVRFTVSGSGVFGKLAALL 266
 QY 124 IPPNFLDMISPAQITMLPHLIVDVRTLEPIPTLPDVRNVYHFNQPPQMRRLVAMLY 183
 DB 267 VPPG--VEPIESVSLQYPHVLFDAQTPEVFTTIPDKIRKLFHSMDETD--TKLVINPY 323
 QY 184 TPLRNSGGDDVFTVSCRVLTRPTDPEFIVLVPP 218
 DB 324 ----ENGV-ENKTCISITVETPSADFTFALLKPP 353

RESULT 10
 ID POLG_FMDV1 STANDARD; PRT; 2333 AA.
 AC P03306; Q64768; Q84750; Q84751; Q84752; Q84753; Q84754; Q84760;
 AC Q84761; Q84762; Q84763; Q84764; Q84765; Q84766; Q84767; Q84768;
 AC Q84769; Q89824;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat
 proteins VP1 TO VP4; Core protein P52; Genome-linked proteins VP61 TO
 VP63; Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed
 RNA polymerase P56A (EC 2.7.7.48)]
 OS Foot-and-mouth disease virus [strain A10-61] (Aphthovirus A) (FMDV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Aphthovirus.
 OX NCBI_TaxID=12112;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84169547; PubMed=6324120;
 RA Carroll A.R., Rowlands D.J., Clarke B.E.;
 RT "The complete nucleotide sequence of the RNA coding for the primary
 translation product of foot and mouth disease virus.";
 RL Nucleic Acids Res. 12:2461-2472(1984).
 RN [2]
 RP SEQUENCE OF 115-1048 FROM N.A.
 RX MEDLINE=82211814; PubMed=6282711;
 RA Boothroyd J.C., Harris T.J.R., Rowlands D.J., Lowe P.A.;
 RT "The nucleotide sequence of cDNA coding for the structural proteins
 of foot-and-mouth disease virus.";
 RL Gene 17:153-161(1982).
 CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the
 poliovirus polyprotein. In other picornavirus reactions Glu may be
 substituted for Gln, and Ser or Thr for Gly.
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 (RNA) (N).
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 each of which is composed of one copy each of proteins VP1, VP2,
 VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; V01130; CAA24361.1; -

DR EMBL; X00429; CAA25127.1; -.
 DR MEROPS; C03.008; -.
 DR MEROPS; C28.001; -.
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR004080; FMDVPicoat.
 DR InterPro; IPR008739; Peptidase_C28.
 DR InterPro; IPR001876; RNv.
 DR InterPro; IPR000605; RNA helicase.
 DR InterPro; IPR007095; RNA pol_PS.
 DR InterPro; IPR001205; RNA pol_P3D.
 DR InterPro; IPR007094; RNA pol_P5vir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF03408; Peptidase_C28; 1.
 DR Pfam; PF00073; rhv; 3.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICVIRUSNS.
 DR PRINTS; PR01542; FMDVPICOAT.
 KW Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;
 KW Transferase; Hydrolase; Thiol protease; Nonstructural protein;
 KW Myristate; Lipoprotein.
 FT CHAIN 1 201 NONSTRUCTURAL PROTEIN P20A.
 FT CHAIN 202 286 COAT PROTEIN VP4.
 FT CHAIN 287 504 COAT PROTEIN VP2.
 FT CHAIN 505 725 COAT PROTEIN VP3.
 FT CHAIN 726 937 COAT PROTEIN VP1.
 FT CHAIN 938 1578 CORE PROTEIN P52.
 FT CHAIN 1579 1601 GENOME-LINKED PROTEIN VP61.
 FT CHAIN 1602 1625 GENOME-LINKED PROTEIN VP62.
 FT CHAIN 1626 1849 GENOME-LINKED PROTEIN VP63.
 FT CHAIN 1850 1863 PROTEASE P20B.
 FT CHAIN 1864 2333 RNA-DIRECTED RNA POLYMERASE P56A.
 FT CHAIN 202 202 N-mvriestoyl glycine (in host).
 FT LIPID 396 396 S -> C (IN REF. 2).
 FT CONFLICT 632 632 P -> L (IN REF. 2).
 FT CONFLICT 632 632
 SQ SEQUENCE 2333 AA; 259645 MW; 4FC667DCC521BC60 CRC64;

Query Match 5.1%; Score 145.5; DB 1; Length 2333;
 Best Local Similarity 20.8%; Pred. No. 0.02;
 Matches 92; Conservative 47; Mismatches 203; Indels 101; Gaps 13;

QY 70 NSPGHILNLELGPDLNPNYLAHLSRMNGYAGGVEVQVLLAGNAFTAGKILFAALPPNPL 129
 DB 358 DKPFGYLTKELEPTDHHGVFGHLVDSYAYMRNGMDVEVSAGVQNGGCLLVAMVPEWKA 417
 QY 130 VDMISPAQITMLPHLIVDVRTLEPIPTLPDVRNVYHFNQPPQMRRLVAMLYTLRSN 189
 DB 418 FDTREKYQLTLFPHQFI-----SPRTNMTAHTVYVLGV 451
 QY 190 GSGDDVFTVSCRVLTRPTDPEFIVLVPPSVESKTPFTLPILTISELTNSRFPPIBOL 249
 DB 452 NRYDQY-----KQHKPWLTVVWLSPLTVSTNAPQIKV 485
 QY 250 YT-APNENNVQCONGRCTDDELQGTOLLSSAVCFLOQRTVADNGNDQNLQLTY 307
 DB 486 YANIAPTIVYHVA-----GELPSKEGIFPVACADGCGGLVTDPTADPVYGVKNV 535
 QY 308 PNGASY-----DPTDEVAPLGTQDFSGMLYGVLTQDNV-----NVSTGEAKNAG 353
 DB 536 PKNTYNGRFTNLLDVAEACPTFLRFDD--GKPYVVTTRADDTRLAKFDVSLA--AKHMSN 592
 QY 354 IYVISTTSKFTPKGISIGLHSSI---TEHVHPNQSRFTPVGVAVDVNTFPQQWVLPHVAG 410
 DB 593 TYLSGIAQYTYQYSTINLHFMFTGSTDASKARYWAYIYPPGVETPDTPPEAAHCHIAEM 652
 QY 411 SLALNTNLAAPVAPTPGEGQLLFFRSRVPVCGVQGGQDAFIDCLLPQBWNNHYQEAAPS 470
 DB 653 DTGLNSKFTFSI-PYVSAADYAYTASDTAETTNVQG-----WVCVYQITHGKA 699
 QY 471 QADVALLIYVNPDTGRTLFEAKL 493
 DB 700 ENDTLL-----VSASAGKO-FEURL 718

DR InterPro; IPR007094; RNA_pol_PSVir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF05408; Peptidase_C28; 1.
DR Pfam; PF00073; rhv; 3.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICVIRUSNS.
DR PRINTS; PR01542; FMDVP1COAT.
DR Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;
KW Transferase; Hydrolase; Thiol protease; Nonstructural protein;
KW Myristate; Lipoprotein.
KW CHAIN 1 200 NONSTRUCTURAL PROTEIN P20A.
FT CHAIN 201 285 COAT PROTEIN VP4.
FT CHAIN 286 503 COAT PROTEIN VP2.
FT CHAIN 504 723 COAT PROTEIN VP3.
FT CHAIN 724 937 COAT PROTEIN VP1.
FT CHAIN 938 953 CORE PROTEIN X.
FT CHAIN 954 1107 CORE PROTEIN P4.
FT CHAIN 1108 1425 CORE PROTEIN P19.
FT CHAIN 1426 1578 GENOME-LINKED PROTEIN VPGL.
FT CHAIN 1579 1601 GENOME-LINKED PROTEIN VPGL.
FT CHAIN 1602 1625 GENOME-LINKED PROTEIN VPGL.
FT CHAIN 1626 1649 GENOME-LINKED PROTEIN VPGL.
FT CHAIN 1650 1862 PROTEASE.
FT CHAIN 1863 2332 RNA-DIRECTED RNA POLYMERASE.
FT CHAIN 201 201 N-myristoyl glycine (in host).
SQ SEQUENCE 2332 AA; 259408 MW; EE77DA739CBEDC6A CRC64;
Query Match 4.5%; Score 130; DB 1; Length 2332;
Best Local Similarity 22.4%; Pred. No. 0.28;
Matches 76; Conservative 31; Mismatches 149; Indels 84; Gaps 10;
QY 79 LELGPDNLNLYLAHLRMYNGYAGGVQVLLAGNAFTAGKILPAALPPHFLVDMISPAQI 138
DB 366 LELFTDHHGVFGLVDSYAYMRNGWDVEVSAGVQNGGCLLVAMVPKMTFTDREYQL 425
QY 139 TMLPHLIVDRTLEPIMTPLPDVNRVYHFNQPPRMRLVAMLYPLRLNSGSDVDFTV 198
DB 426 TLFPHQFI-----SPRTNTHAITYPLGVNRYDQY--- 456
QY 199 SCRVLTRPTDFEYILVPPSVESKTPFTLPILTISELNSRFPPIEOLYT--APNET 256
DB 457 -----KKHKPWTLLVIMVLSPLTVSNTAATQIKVYANIAITYV 493
QY 257 NVVQCGNRCITLDELQGTQLSSAVCFQGRV-----ADNGDNWQNLQLQTYPNG 310
DB 494 HVA-----GELPSKVGIFPVACSDGYGLVTTDPKTADPVYKYNPPKTYNPRR 543
QY 311 AS--YDPTDEVPAPIGTQDFSGMLYGLVLTQDNV-----NVSTGEAKNAKGIYISTSGK 362
DB 544 FTNLLDVAEACPTFLCFDD--GKPYVYVTRTDTRLLAKFDVSLA-AKHMSNTYLSGIAQY 600
QY 363 FTKIGSIGLHSI-----TEHVHPNQOSRTPVGVAVDENTP 399
DB 601 YTYSGSTINLHPMTGSTDSKARYVMVAYIPPGVETPPETP 640
RESULT 12
POLG CX16T
ID POLG CX16T STANDARD; PRT; 2193 AA.
AC Q9QF31;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2
DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein
DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].
OS Coxsackievirus A16 (strain Tainan/5079/98).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OC NCBI_TaxID=231417;
POLG FMDVA
ID POLG FMDVA STANDARD; PRT; 2332 AA.
AC P03308; P03312; Q65038; Q65039; Q65040; Q65041; Q65042; Q65043;
AC Q65044; Q65045; Q65046; Q65047;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat
DE proteins VP1 TO VP4; Core proteins X, P14, P41, P19; Genome-linked
DE proteins VPGL TO VPGL3; Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48)].
OS Foot-and-mouth disease virus (strain A12) (Aphthovirus A) (FMDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Aphthovirus.
OC NCBI_TaxID=12114;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=85211015; PubMed=2987518;
RX Robertson B.H., Grubman M.J., Weddell G.N., Moore D.M., Welsh J.D.,
RA Fischer T., Dowbenko D.J., Yansura D.G., Small B., Kleid D.G.;
RT "Nucleotide and amino acid sequence coding for polypeptides of
RT foot-and-mouth disease virus type A12.";
RL J. Virol. 54:651-660(1985).
RN [2]
RP SEQUENCE OF 1863-2332 FROM N.A.
RX MEDLINE=83225613; PubMed=6305004;
RA Robertson B.H., Morgan D.O., Moore D.M., Grubman M.J., Card J.,
RA Fischer T., Weddell G.N., Dowbenko D.J., Yansura D.G.;
RT "Identification of amino acid and nucleotide sequence of the
RT foot-and-mouth disease virus RNA polymerase.";
RL Virology 126:614-623(1983).
RN [3]
RP SEQUENCE OF 715-955 FROM N.A.
RX MEDLINE=82061853; PubMed=6272395;
RA Kleid D.G., Yansura D.G., Small B., Dowbenko D.J., Moore D.M.,
RA Grubman M.J., McKercher P.D., Morgan D.O., Robertson B.H.,
RA Bachrach H.L.;
RT "Cloned viral protein vaccine for foot-and-mouth disease: responses
RT in cattle and swine.";
RL Science 214:1125-1129(1981).
CC -1- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the
CC poliovirus polypeptide. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA) (N).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M10975; AAA42593.1; -;
DR EMBL; J02187; AAA42670.1; -;
DR MEROPS; C03.008; -;
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR004080; FMDVP1coat.
DR InterPro; IPR008739; Peptidase_C28.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_PS.
DR InterPro; IPR001205; RNA_pol_P3D.

RT Sinorhizobium meliloti pSymA megaplasmid.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888 (2001).
 CC -!- FUNCTION: RECEPTOR FOR THE SIDEROPHORE RHIZOACTIN.
 CC -!- SUBCELLULAR LOCATION: Outer membrane.
 CC -!- SIMILARITY: Belongs to the tonB-dependent receptor family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF110737; AAD09419.1; --
 DR EMBL; AR007312; AAK65923.1; --
 DR PIR; A95420; A95420.
 DR PIR; T46821; T46821.
 DR InterPro; IPR000531; TonB boxC.
 DR Pfam; PF00593; TonB dep Rec; 1.
 DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; FALSE NEG.
 DR PROSITE; PS01156; TONB_DEPENDENT_REC_2; FALSE NEG.
 DR Outer membrane; Iron transport; Transport; TonB box; Signal; Receptor;
 KW Plasmid; Complete proteome.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 746 RHIZOACTIN RECEPTOR.
 FT SITE 40 47 TONB BOX.
 FT SITE 729 746 TONB C-TERMINAL BOX.
 FT SITE 746 AA; 80633 MW; 16A844A025D5B23 RC64;
 SQ SEQUENCE 746 AA; 80633 MW; 16A844A025D5B23 RC64;
 Query Match 4.3%; Score 123.5; DB 1; Length 746;
 Best Local Similarity 19.7%; Pred. No. 0.18;
 Matches 110; Conservative 76; Mismatches 184; Indels 189; Gaps 26;
 Qy 10 PSNDGAASLVPEGINETMPLEPVAGSI--AAPVAGTNIIDPWIRNFVQAPNGEFTVS 67
 Db 91 PASDGARTSGQNLGRPPILIDGVSMNSARLSRQFDALDF-NIERVEVLSGATAIY 149
 Qy 68 PRNSPGEL-LNLEGLFDLPFLYLAH---LSRMVNGVAGV---EVQ 106
 Db 150 GGNATGGIINIITKKGADEPL-HAEVGTGMSGFGAGSQDFDRAAGAVTYNSENWDAR 208
 Qy 107 VLLAGN---AF--TAGKILFAAIPP---NFLVDMISPAITMLPHLIVDVRTLE--- 152
 Db 209 LSIAGNRTGAFYDGSGLTLLIDITQTSTAFNERIDLMG-----SIGQIDDDRVEFSQG 263
 Qy 153 -----PIMTLPEDVRNVF-----YHFNQPPQNRVLVAMLYT----- 184
 Db 264 YFDSKQSDGYLYGPPFAALAD-PSLPETRSGYSEDFNPQTRRSLNVTYTDNDVFGQQ 322
 Qy 185 -----PLRNSGSD-----DVFTVSCVLRTPPTDPFEIYLP 217
 Db 323 LLLQGSYRTERIKFHPFPASGNSETGPYFGSSQTDYIGIRAAALVAEPTALKITYGID 382
 Qy 218 PSVESKT-----KPFTLPILITSELNSRFPPIPIEQLYTAPN 254
 Db 383 ADMDSFTARQNTFDMVAACQSGGLDFNTIGKTGLYPSIDVSTVAGF-----AEASYEATD 437
 Qy 255 ETNVVQCQNGRCLTDLGELQ-----GTTQLLSANVCFLQGR-TVADNGDNWDQN 301
 Db 438 -----RLTLNGGVRYQFNVTEVSDFIGAAQ-----QVAILOGRATSADTTPGSEVN 483
 Qy 302 LIQLTYPCASVDPTD-----EVRAP-----LGTQDFSGMLYGLVLTQDNVNS 344
 Db 484 YDAALFSAGATVQLTNTQOYVAFNSQGFELPPAKYGYGNYVFSFGHYTLVNSVNVGDS 543
 Qy 345 TGAAKNAKGIYISTTSKFTPKIGSTGLHSITEH-VHPNQQRFTPV----- 390
 Db 544 ALBAIKTNSFEIGYRLDDGTFTNLEATYAAYSLSDRSINLNRSLAVEIIDRRERYVYIEGK 603
 Qy 391 -GVAVDENTPFQOWLPY 408
 Db 604 AGVKLDHG--FDVGLVGHW 620

RESULT 14

POLG HE701 STANDARD; PRT; 2194 AA.
 ID POLG HE701 STANDARD; PRT; 2194 AA.
 AC P32537;
 DT 01-OCT-1993 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 42, Last annotation update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polypeptide (P1C); Coat protein VP4 (P1A); Coat protein VP2
 DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein
 DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-
 DE linked protein VPG (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)
 DE (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48).
 OS Human enterovirus 70 (strain J670/71) (EV 70).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Enterovirus.
 OC NCBI_TaxID=31915;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91037960; PubMed=2172447;
 RA Ryan M.D., Jenkins O., Hughes P.J., Brown A., Knowles N.J., Booth D.,
 RA Minor P.D., Almond J.W.;
 RT "The complete nucleotide sequence of enterovirus type 70:
 RT relationships with other members of the picornaviridae.";
 RL J. Gen. Virol. 71:2291-2299 (1990).
 CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
 CC O/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
 CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-[Gly bond in the
 CC poliovirus polyprotein. In other picornavirus reactions Glu may be
 CC substituted for Gln, and Ser or Thr for Gly.
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA) (N).
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D00820; BAA18891.1; --
 DR PIR; A36253; GNNYE7.
 DR HSP; P03300; IPOV.
 DR MEROPS; C03.UBA; --
 DR MEROPS; C03.UBA; --
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR004004; Calici pol hel.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR001199; pept 3C picorn.
 DR InterPro; IPR000081; peptidase C3.
 DR InterPro; IPR003338; Pico_P1A.
 DR InterPro; IPR002527; Pico_P2B.
 DR InterPro; IPR001676; Rhv.
 DR InterPro; IPR006005; RNA helicase.
 DR InterPro; IPR007095; RNA pol DS PS.
 DR InterPro; IPR001205; RNA pol P35.
 DR InterPro; IPR007094; RNA pol PSvir.
 DR InterPro; IPR008975; Viral cap coat.
 DR Pfam; PF00548; Cys-protease-3C; 1.
 DR Pfam; PF02226; Pico_P1A; 1.
 DR Pfam; PF00947; Pico_P2B; 1.
 DR Pfam; PF01552; Pico_P2B; 1.
 DR Pfam; PF00073; rhv; 3.
 DR Pfam; PF00680; RNA dep RNA pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICVIRUSNS.

```

DR ProDom; PD001125; Cys_protease_3C; 1.
DR ProDom; PD001306; Pico_P2A; 1.
DR ProDom; PD001274; Pico_P2B; 1.
DR SMART; SM00382; AAA; 1.
KW Polypeptide; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
KW Lipoprotein. 2 69 COAT PROTEIN VP4.
FT CHAIN 70 319 COAT PROTEIN VP2.
FT CHAIN 320 561 COAT PROTEIN VP3.
FT CHAIN 324 565 COAT PROTEIN VP1.
FT CHAIN 872 1014 CORE PROTEIN P2A.
FT CHAIN 1015 1113 CORE PROTEIN P2B.
FT CHAIN 1114 1443 CORE PROTEIN P2C.
FT CHAIN 1444 1532 CORE PROTEIN P3A.
FT CHAIN 1533 1554 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1555 1737 PICORNAIN 3C.
FT CHAIN 1738 2194 RNA-DIRECTED RNA POLYMERASE P3D.
FT LIPID 2 2 N-myristoyl glycine (in host) (by
FT ACT SITE 1701 1701 PROTEASE (POTENTIAL).
FT ACT SITE 1715 1715 PROTEASE (POTENTIAL).
SQ SEQUENCE 2194 AA; 244590 MW; 15DBAE96EE06673C CRC64;

Query Match 4.1%; Score 118; DB 1; Length 2194;
Best Local Similarity 21.9%; Pred. No. 2;
Matches 57; Conservative 39; Mismatches 114; Indels 50; Gaps 10;

QY 48 IDPWRTNFWQPNQ-----EFTVSPNSGTEILLNLEL-----GPDLPYLAHLGRMYNG 98
DQ 368 IESMEINNNNDASGVERLURVOISQSDMDQLFNIPDLIQEGPLRWLGNISRYTH 427
QY 99 YAGGVEVQVLLAGNAFTAGKILFAAPFNFLVDMISPA---QITMLPHLIVDVRLTEPIM 155
DQ 428 WSGSLEMTFMFCGSPMTTGLIICVTPG---GSSPTDRMQAMLAHVWDFGLQSSIT 483
QY 156 TPLDVRNVFYH-FNNQOPQRMVAMLYTPLRN-----GSGDVFVTVSCRV----- 202
DQ 484 IIPWISGSHYRMFTDAKAINANVGYVTCFMQTLNAPVGAADQCYIVGMVAKKDFNL 543
QY 203 -LTRPTDPF-----EFYLVPPSVESKTKP--FTLPILTISE--LTNSR 241
DQ 544 RLRDSDPDIGQAIILPEQAATQIGEVKTVANTVESEIKAEGLVPSLNAVETGATNT 603
QY 242 FPIPIEQLYTAPNETNVMVQC 261
DQ 604 EPEEALQRTVNLNMGHTAEC 623

RESULT 15
POLG_CX16G
ID POLG_CX16G STANDARD; PRT; 2193 AA.
AC Q65900.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polypeptide [Contains: Coat protein VP4 (PIA); Coat protein VP2
DE (PIB); Coat protein VP3 (PIC); Coat protein VP1 (PID); Core protein
DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-
DE linked protein VPG (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].
OS Coxsackievirus A16 (strain G-10).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OC NCBI_TaxID=69159;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94303216; PubMed=8030260;
RT Povty T., Hyslop T., Horsnell C., Kinnunen L., Hovi T., Stanway G.;
RT "Molecular analysis of coxsackievirus A16 reveals a new genetic group
RT of enteroviruses.";
RL Virology 202:982-987(1994).
CC -!- FUNCTION: It is thought that the P2C protein attaches to vesicular

```

```

CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the
CC poliovirus polypeptide. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA] (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; U05876; AAA50478.1; -.
CC HSP; P03300; IPOV.
CC MEROPS; C03.022; -.
CC DR InterPro; IPR003593; AAA_ATPase.
CC DR InterPro; IPR004004; Calici_pol_hel.
CC DR InterPro; IPR009003; Cys_Ser_trypsin.
CC DR InterPro; IPR001199; Pept_3C_picorn.
CC DR InterPro; IPR000081; Peptidase_C3.
CC DR InterPro; IPR003138; Pico_P1A.
CC DR InterPro; IPR002527; Pico_P2B.
CC DR InterPro; IPR001676; Rnv.
CC DR InterPro; IPR000605; RNA_helicase.
CC DR InterPro; IPR007095; RNA_pol_DS_P8.
CC DR InterPro; IPR001205; RNA_pol_P3D.
CC DR InterPro; IPR007094; RNA_pol_PSVir.
CC DR InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00548; Cys-protease-3C; 1.
CC Pfam; PF02226; Pico_P1A; 1.
CC Pfam; PF00947; Pico_P2A; 1.
CC Pfam; PF01552; Pico_P2B; 1.
CC Pfam; PF00073; thv; 3.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICVIRUSNS.
CC ProDom; PD001125; Cys_protease_3C; 1.
CC ProDom; PD001306; Pico_P2A; 1.
CC ProDom; PD001274; Pico_P2B; 1.
CC SMART; SM00382; AAA; 1.
CC KW RNA-directed RNA polymerase; Core protein; Core protein; Transferase;
KW Lipoprotein.
KW CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 323 COAT PROTEIN VP2.
FT CHAIN 324 565 COAT PROTEIN VP3.
FT CHAIN 566 862 COAT PROTEIN VP1.
FT CHAIN 863 1012 CORE PROTEIN P2A.
FT CHAIN 1013 1111 CORE PROTEIN P2B.
FT CHAIN 1112 1440 CORE PROTEIN P2C.
FT CHAIN 1441 1526 CORE PROTEIN P3A.
FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1549 1731 PICORNAIN 3C.
FT CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE.
FT LIPID 2 2 N-myristoyl glycine (in host) (by
FT ACT SITE 1695 1695 PROTEASE (POTENTIAL).
FT ACT SITE 1709 1709 PROTEASE (POTENTIAL).
SQ SEQUENCE 2193 AA; 243209 MW; 04B3BC572A76E38 CRC64;

Query Match 4.1%; Score 116.5; DB 1; Length 2193;
Best Local Similarity 20.9%; Pred. No. 2.6;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: June 1, 2004, 13:35:27 ; Search time 31.556 Seconds
(without alignments)
5419.275 Million cell updates/sec

Title: US-09-926-799-9
Perfect score: 2863
Sequence: 1 MKMASNDAPNDGASLVP.....VNQFSLAPMGTCNGRRRIQ 542

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriaph.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2863	100.0	542	12 Q8JW45	Q8JW45 norwalk-lik
2	2822	98.6	542	12 Q8JW45	Q8JW45 norwalk-lik
3	2049.5	71.6	535	12 Q91115	Q91115 human calic
4	2048.5	71.6	535	12 Q9QMK6	Q9QMK6 chitta viru
5	2047.5	71.5	535	12 Q917X7	Q917X7 norwalk-lik
6	2046.5	71.5	535	12 Q915C6	Q915C6 human calic
7	2028.5	70.9	535	12 Q8V784	Q8V784 norwalk vir
8	2020.5	70.6	535	12 Q88104	Q88104 hawaii cali
9	2016	70.4	548	12 Q915C7	Q915C7 human calic
10	2015.5	70.4	535	12 Q917Y9	Q917Y9 norwalk-lik
11	2010.5	70.2	535	12 Q917Y3	Q917Y3 norwalk-lik
12	2007.5	70.1	535	12 Q917X4	Q917X4 human calic
13	2006.5	70.1	535	12 Q916E4	Q916E4 human calic
14	2003.5	70.0	535	12 Q917Y0	Q917Y0 norwalk-lik
15	2002.5	69.9	535	12 Q916E6	Q916E6 human calic
16	1994.5	69.7	535	12 Q91V39	Q91V39 human calic

17	1988.5	69.5	535	12 Q91V51	Q91V51 human calic
18	1980	69.2	542	12 Q66293	Q66293 calicivirus
19	1973	68.9	542	12 Q12305	Q12305 snow mounta
20	1973	68.9	542	12 Q8ORD6	Q8ORD6 snow mounta
21	1970	68.8	542	12 Q96877	Q96877 snow mounta
22	1966	68.7	542	12 Q913B6	Q913B6 human calic
23	1943	67.9	548	12 Q66296	Q66296 calicivirus
24	1942	67.8	548	12 Q8V0P5	Q8V0P5 human calic
25	1942	67.8	548	12 Q8V0P4	Q8V0P4 human calic
26	1935.5	67.6	547	12 Q8V775	Q8V775 norwalk vir
27	1934.5	67.6	547	12 Q918A7	Q918A7 norwalk-lik
28	1933.5	67.5	547	12 Q8V773	Q8V773 norwalk vir
29	1928	67.3	540	12 Q917X1	Q917X1 norwalk-lik
30	1928	67.3	548	12 Q8V768	Q8V768 norwalk vir
31	1928	67.3	548	12 Q918A1	Q918A1 norwalk-lik
32	1928	67.3	548	12 Q8BC87	Q8BC87 human calic
33	1924	67.2	540	12 Q917W8	Q917W8 norwalk-lik
34	1923	67.2	548	12 Q916E5	Q916E5 human calic
35	1922	67.1	548	12 Q83880	Q83880 norwalk vir
36	1921	67.1	548	12 Q88291	Q88291 small round
37	1921	67.1	548	12 Q80FL1	Q80FL1 human calic
38	1920	67.1	548	12 Q80FK8	Q80FK8 human calic
39	1920	67.1	548	12 Q80FK5	Q80FK5 human calic
40	1919	67.0	548	12 Q80FK9	Q80FK9 human calic
41	1918	67.0	538	12 Q91165	Q91165 human calic
42	1918	67.0	548	12 Q91725	Q91725 norwalk-lik
43	1916	66.9	548	12 Q9PYA7	Q9PYA7 human calic
44	1916	66.9	548	12 Q8BC90	Q8BC90 human calic
45	1916	66.9	548	12 Q80FK4	Q80FK4 human calic

ALIGNMENTS

RESULT 1
Q8JW45 PRELIMINARY; PRT; 542 AA.
ID Q8JW45
AC Q8JW45;
DC 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Capsid protein.
DE CAPSID.
GN Norwalk-like virus.
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OC NCBI_TaxID=95340;
OX NCBI_TaxID=95340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kashiwa47;
RA Natori K., Takeda N.;
RT "Genetic and antigenic relationship among Norwalk-like viruses.";
RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB078334; BAC05515.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 542 AA; 59149 MW; 5C28719C65F6C667 CRC64;

Query Match 100.0%; Score 2863; DB 12; Length 542;
Best Local Similarity 100.0%; Pred. No. 6.9e-217;
Matches 542; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKMASNDAPNDGASLVP...VNQFSLAPMGTCNGRRRIQ 542
Db 1 MKMASNDAPNDGASLVP...VNQFSLAPMGTCNGRRRIQ 542
QY 61 NGEFTVSPRNSFGTEILLNLELGPDLNLYLAHLRMYNGYAGGVEVQVLLAGNAFTAGKIL 120

Db	61	NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLRMVNGYAGGVEQVLLAGNAFTAGKIL	120	QY	121	FAAIPNPNFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNNOQPRMRLVA	180
QY	121	FAAIPNPNFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNNOQPRMRLVA	180	Db	121	FAAIPNPNFVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNNOQPRMRLVA	180
Db	121	FAAIPNPNFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNNOQPRMRLVA	180	QY	181	MLYTPLRNSGSGDDVFTVSCRVLTPTDPEFIYLVPPSVESKTKPFTLPILTISELTNS	240
QY	181	MLYTPLRNSGSGDDVFTVSCRVLTPTDPEFIYLVPPSVESKTKPFTLPILTISELTNS	240	Db	181	MLYTPLRNSGSGDDVFTVSCRVLTPTDPEFIYLVPPSVESKTKPFTLPILTISELTNS	240
QY	181	MLYTPLRNSGSGDDVFTVSCRVLTPTDPEFIYLVPPSVESKTKPFTLPILTISELTNS	240	QY	241	RPPIPIEQIYTAAPNETNVVQCGRCITLDGELQGTQTLSSAVCFLOGRTVADRGDNWDQ	300
Db	181	MLYTPLRNSGSGDDVFTVSCRVLTPTDPEFIYLVPPSVESKTKPFTLPILTISELTNS	240	Db	241	RPPIPIEQIYTAAPNETNVVQCGRCITLDGELQGTQTLSSAVCFLOGRTVADRGDNWDQ	300
QY	241	RPPIPIEQIYTAAPNETNVVQCGRCITLDGELQGTQTLSSAVCFLOGRTVADRGDNWDQ	300	QY	301	NULLQTYPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKIYISTTS	360
Db	241	RPPIPIEQIYTAAPNETNVVQCGRCITLDGELQGTQTLSSAVCFLOGRTVADRGDNWDQ	300	Db	301	NULLQTYPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKIYISTTS	360
QY	301	NULLQTYPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKIYISTTS	360	QY	361	GKFTPKIGSIGLHSITEHVHPNQOSRFTPVGVAVDENTPFQOWVLPHYAGSLALNTNLAP	420
Db	301	NULLQTYPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKIYISTTS	360	Db	361	GKFTPKIGSIGLHSITEHVHPNQOSRFTPVGVAVDENTPFQOWVLPHYAGSLALNTNLAP	420
QY	361	GKFTPKIGSIGLHSITEHVHPNQOSRFTPVGVAVDENTPFQOWVLPHYAGSLALNTNLAP	420	QY	421	AVAPTPFGSQLLFFRSRVPVCVQLOGQDAFIDCLLPQEWVNHFYQEAAPSQADVALIRYV	480
Db	361	GKFTPKIGSIGLHSITEHVHPNQOSRFTPVGVAVDENTPFQOWVLPHYAGSLALNTNLAP	420	Db	421	AVAPTPFGSQLLFFRSRVPVCVQLOGQDAFIDCLLPQEWVNHFYQEAAPSQADVALIRYV	480
QY	421	AVAPTPFGSQLLFFRSRVPVCVQLOGQDAFIDCLLPQEWVNHFYQEAAPSQADVALIRYV	480	QY	481	NPDGTGRTLFEAKLHRSGLFTVSHGTGAYPLVVPNGHFRFDSWMNQFYSLAPMGTGNRRR	540
Db	421	AVAPTPFGSQLLFFRSRVPVCVQLOGQDAFIDCLLPQEWVNHFYQEAAPSQADVALIRYV	480	Db	481	NPDGTGRTLFEAKLHRSGLFTVSHGTGAYPLVVPNGHFRFDSWMNQFYSLAPMGTGNRRR	540
QY	541	IQ 542		QY	541	IQ 542	
Db	541	IQ 542		Db	541	IQ 542	
RESULT 3							
Q91115	Q91115	PRELIMINARY;	PRT;	535	AA.		
AC	Q91115;						
DT	01-DEC-2001 (TrEMBLrel. 19, Created)						
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)						
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)						
DE	Major capsid protein.						
OS	Human calicivirus NLV/Schwerin 003/00/DE.						
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;						
OC	Norovirus.						
OX	NCBI_TaxID=168852;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RA	Kuenkel U., Schreier E.;						
RT	"Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany."						
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.						
DR	EMBL; AF397905; AAK92128.1; --						
DR	InterPro; IPR004005; Calici_coat.						
DR	InterPro; IPR008975; Viral_cap_coat.						
DR	Pfam; PF00915; Calici_coat; 1.						
SQ	SEQUENCE 535 AA; 58520 MW; 2597BBD3F3D5FCC1 CRC64;						
Query Match 71.6%; Score 2049.5; DB 12; Length 535;							
Best Local Similarity 71.7%; Pred. No. 8.1e-153;							
Matches 392; Conservative 65; Mismatches 73; Indels 17; Gaps 7;							
QY	1	MKMASNDAAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQTNIIDPWIRNFVQAP	60	QY	1	MKMASNDAAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQTNIIDPWIRNFVQAP	60
Db	1	MKMASNDAAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQTNIIDPWIRNFVQAP	60	Db	1	MKMASNDAAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQTNIIDPWIRNFVQAP	60
QY	61	NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLRMVNGYAGGVEQVLLAGNAFTAGKIL	120	QY	61	NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLRMVNGYAGGVEQVLLAGNAFTAGKIL	120
Db	61	NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLRMVNGYAGGVEQVLLAGNAFTAGKIL	120	Db	61	NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLRMVNGYAGGVEQVLLAGNAFTAGKIL	120
QY	121	FAAIPNPNFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNNOQPRMRLVA	180	QY	121	FAAIPNPNFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNNOQPRMRLVA	180
Db	121	FAAIPNPNFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNNOQPRMRLVA	180	Db	121	FAAIPNPNFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNNOQPRMRLVA	180

QY 181 MLYTPLRNSGDDVFTVSCRVLTTRPTDPDFEYILVPPSVESKTKPFTLPILTISLTNS 240
 DB 181 MLYTPLRNSGDDVFTVSCRVLTTRPTDPDFEYILVPPSVESKTKPFTLPILTISLTNS 240
 QY 241 RPIPIEOLYTAPNETNVVQCGRCCTLDGELQGTTLSSAVCFLOGR---TVADNGDN 297
 DB 241 RPIPIEOLYTAPNETNVVQCGRCCTLDGELQGTTLSSAVCFLOGR---TVADNGDN 297
 QY 298 WDQNLQLTYPNGASYDPTDEVPAPPLGTQDFSGMLYGVLTQ--DNVNVSTGEAKNAKIYI 356
 DB 298 WDQNLQLTYPNGASYDPTDEVPAPPLGTQDFSGMLYGVLTQ--DNVNVSTGEAKNAKIYI 356
 QY 301 WN---MQVTNINGTPDPTDVPAPPLGTQDFSGMLYGVLTQ--DNVNVSTGEAKNAKIYI 352
 DB 301 WN---MQVTNINGTPDPTDVPAPPLGTQDFSGMLYGVLTQ--DNVNVSTGEAKNAKIYI 352
 QY 357 STTSCKFTPKIGSLGHSITE-IHVPNQOSRFTPVGVAVDENTPQQWVLPYAGSLAIN 415
 DB 357 STTSCKFTPKIGSLGHSITE-IHVPNQOSRFTPVGVAVDENTPQQWVLPYAGSLAIN 415
 QY 416 TNLPAVAPTFFGEOLLEFRSRVPCVQGLQGDADFDCLLPQEWVNHFFVQEAAPQADVA 475
 DB 416 TNLPAVAPTFFGEOLLEFRSRVPCVQGLQGDADFDCLLPQEWVNHFFVQEAAPQADVA 475
 QY 476 LIRVYNPDTGRTLFEAKLHRSFTVSHGTGAYPLVVPNGHFRFDSWVNFYSLAPMG 535
 DB 476 LIRVYNPDTGRTLFEAKLHRSFTVSHGTGAYPLVVPNGHFRFDSWVNFYSLAPMG 535
 QY 536 NGRRIQ 542
 DB 536 NGRRIQ 542
 QY 529 NGRRRVQ 535
 DB 529 NGRRRVQ 535

RESULT 4

Q9QMK6 PRELIMINARY; PRT; 535 AA.
 AC Q9QMK6;
 DT 01-MAY-2000 (TremBLrel. 13, Created)
 DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
 DE Capsid protein.
 OS Chittia virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 OX NCBI_TaxID=104901;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Aichi76-96;
 RX MEDLINE=20471382; PubMed=11021399;
 RA Kobayashi S., Sakae K., Suzuki Y., Ieshiko H., Kamata K., Suzuki K.,
 RA Natori K., Miyamura T., Takeda N.;
 RT "Expression of recombinant capsid proteins of chittia virus, a
 RT genogroup II Norwalk virus, and development of an ELISA to detect the
 RT viral antigen.";
 RL Microbiol. Immunol. 44:687-693(2000).
 DR EMBL; AB032758; BAA84716.1; -;
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 SQ SEQUENCE 535 AA; 58504 MW; 5CCEBC91BFC7B8B CRC64;

Query Match 71.6%; Score 2048.5; DB 12; Length 535;
 Best Local Similarity 71.7%; Pred. No. 9.8e-153;
 Matches 392; Conservative 64; Mismatches 74; Indels 17; Gaps 7;

QY 1 MKMASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQNTIIDPWIRTFVQAP 60
 DB 1 MKMASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQNTIIDPWIRTFVQAP 60
 QY 61 NGEFTVSPNSGCEILLNLEIGDPLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKIL 120
 DB 61 NGEFTVSPNSGCEILLNLEIGDPLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKIL 120
 QY 121 FAATPPLVDNISPAQITMLPHLVDVTRLEPIMTPLPDVENVFHFNNQPPRLVA 180
 DB 121 FAATPPLVDNISPAQITMLPHLVDVTRLEPIMTPLPDVENVFHFNNQPPRLVA 180

QY 181 MLYTPLRNSGDDVFTVSCRVLTTRPTDPDFEYILVPPSVESKTKPFTLPILTISLTNS 240
 DB 181 MLYTPLRNSGDDVFTVSCRVLTTRPTDPDFEYILVPPSVESKTKPFTLPILTISLTNS 240
 QY 241 RPIPIEOLYTAPNETNVVQCGRCCTLDGELQGTTLSSAVCFLOGR---TVADNGDN 297
 DB 241 RPIPIEOLYTAPNETNVVQCGRCCTLDGELQGTTLSSAVCFLOGR---TVADNGDN 297
 QY 298 WDQNLQLTYPNGASYDPTDEVPAPPLGTQDFSGMLYGVLTQ--DNVNVSTGEAKNAKIYI 356
 DB 298 WDQNLQLTYPNGASYDPTDEVPAPPLGTQDFSGMLYGVLTQ--DNVNVSTGEAKNAKIYI 356
 QY 301 WN---MQVTNINGTPDPTDVPAPPLGTQDFSGMLYGVLTQ--DNVNVSTGEAKNAKIYI 352
 DB 301 WN---MQVTNINGTPDPTDVPAPPLGTQDFSGMLYGVLTQ--DNVNVSTGEAKNAKIYI 352
 QY 357 STTSCKFTPKIGSLGHSITE-IHVPNQOSRFTPVGVAVDENTPQQWVLPYAGSLAIN 415
 DB 357 STTSCKFTPKIGSLGHSITE-IHVPNQOSRFTPVGVAVDENTPQQWVLPYAGSLAIN 415
 QY 416 TNLPAVAPTFFGEOLLEFRSRVPCVQGLQGDADFDCLLPQEWVNHFFVQEAAPQADVA 475
 DB 416 TNLPAVAPTFFGEOLLEFRSRVPCVQGLQGDADFDCLLPQEWVNHFFVQEAAPQADVA 475
 QY 476 LIRVYNPDTGRTLFEAKLHRSFTVSHGTGAYPLVVPNGHFRFDSWVNFYSLAPMG 535
 DB 476 LIRVYNPDTGRTLFEAKLHRSFTVSHGTGAYPLVVPNGHFRFDSWVNFYSLAPMG 535
 QY 536 NGRRIQ 542
 DB 536 NGRRIQ 542
 QY 529 NGRRRVQ 535
 DB 529 NGRRRVQ 535

RESULT 5
 Q917X7 PRELIMINARY; PRT; 535 AA.
 AC Q917X7;
 DT 01-DEC-2001 (TremBLrel. 19, Created)
 DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
 DE Capsid protein.
 OS Norwalk-like virus NLV/Honolulu/314/1994/US.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 OX NCBI_TaxID=171852;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Hu/NLV/Honolulu/314/1994/US;
 RX MEDLINE=97193806; PubMed=9041391;
 RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
 RT "A one-tube method of reverse transcription-PCR to efficiently amplify
 RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
 RT of small round-structured viruses (Norwalk-like viruses).";
 RL J. Clin. Microbiol. 35:570-577(1997).
 RN [2]
 RP SEQUENCE OF 100-192 FROM N.A.
 RC STRAIN=Hu/NLV/Honolulu/314/1994/US;
 RX MEDLINE=98071277; PubMed=9407386;
 RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
 RA Seto Y., Monroe S.S., Glass R.I.;
 RT "Correlation of patient immune responses with genetically
 RT characterized small round-structured viruses involved in outbreaks of
 RT nonbacterial acute gastroenteritis in the United States, 1990 to
 RT 1995.";
 RL J. Med. Virol. 53:372-383(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Hu/NLV/Honolulu/314/1994/US;
 RX MEDLINE=20266071; PubMed=10804147;
 RA Ando T., Noel J.S., Fankhauser R.L.;
 RT "Genetic classification of 'Norwalk-like viruses.'";
 RL J. Infect. Dis. 181:S336-S348(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Hu/NLV/Honolulu/314/1994/US;
 RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;

```
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1. D839C359F86DC1F9 CRC64;
SQ SEQUENCE 535 AA; 58505 MW; 71.5% Score 2046.5; DB 12; Length 535;
Query Match 71.5%; Score 2046.5; DB 12; Length 535;
Best Local Similarity 71.7%; Pred. No. 1.4e-152;
Matches 392; Conservative 64; Mismatches 74; Indels 17; Gaps 7;
1 MKWASNDAAPSNDGASLVPEGINETMPLEPVAGASIAAPVAGQTNIIDPWIRNFVQAP 60
1 MKWASNDAAPSNDGASLVPEGINETMPLEPVAGASIAAPVAGQTNIIDPWIRNFVQAP 60
61 NGEFTVSPRNSGCEILLNLELGPDLNPLYLAHLSRMVNGYAGGVEQVLLAGNAFTAGKIL 120
61 NGEFTVSPRNSGCEILLNLELGPDLNPLYLAHLSRMVNGYAGGVEQVLLAGNAFTAGKIL 120
121 FFAIPNPLVDMSIPAQITMLPHLIVDRTLEPIMTLPDVRNVPYHFNNOQPMRLVA 180
121 FFAIPNPLVDMSIPAQITMLPHLIVDRTLEPIMTLPDVRNVPYHFNNOQPMRLVA 180
121 FFAVPPHFFLENISPOQITMFEHVIIDVRLPELVLLPLDVRNFFHYNQNEPRMRLVA 180
181 MLYTPLRSNGSGDDVFTVSCRVLTRPTDFEFYIYLVPPSVESKTKPFTLPILTISELTNS 240
181 MLYTPLRSNGSGDDVFTVSCRVLTRPTDFEFYIYLVPPSVESKTKPFTLPILTISELTNS 240
241 RFPPIEQLYTAPNETNVOQNGRCCTLDGELQGTQLLSSAVCFLOGR---TVADNGDN 297
241 RFPPIEQLYTAPNETNVOQNGRCCTLDGELQGTQLLSSAVCFLOGR---TVADNGDN 297
298 WDQNLQLQTYPNGASVDPDEVPAPLGTQDFSGMLYGVLTQ--DNVNVSTGEAKNAKIYI 356
298 WDQNLQLQTYPNGASVDPDEVPAPLGTQDFSGMLYGVLTQ--DNVNVSTGEAKNAKIYI 356
301 WN---MQVTNNGTPEDTDPEDVPAPLGTDFSGRLFGVLSQRDHDN-----ACRSHDAVI 352
301 WN---MQVTNNGTPEDTDPEDVPAPLGTDFSGRLFGVLSQRDHDN-----ACRSHDAVI 352
357 STTSKGKTPKIGSLHSITE-HVHPNQQRFTPVGVAVDENTPFOQWVLPYHAGSLALN 415
357 STTSKGKTPKIGSLHSITE-HVHPNQQRFTPVGVAVDENTPFOQWVLPYHAGSLALN 415
353 ATNSAKTPKLGAIQGTWEEDDVHINQPTKFTPVGLF--EHGDFNQWTLPNYSGALTIN 410
353 ATNSAKTPKLGAIQGTWEEDDVHINQPTKFTPVGLF--EHGDFNQWTLPNYSGALTIN 410
416 TNLAPAVAPTPPGQQLFFRSRVPVCGVQGLQGDADIDCLLPQEWNVHFYQEAAPSQADVA 475
416 TNLAPAVAPTPPGQQLFFRSRVPVCGVQGLQGDADIDCLLPQEWNVHFYQEAAPSQADVA 475
411 MGLAPPVAPTPPGQQLFFRSRVPVCGVQGLQGDADIDCLLPQEWNVHFYQEAAPSQADVA 468
411 MGLAPPVAPTPPGQQLFFRSRVPVCGVQGLQGDADIDCLLPQEWNVHFYQEAAPSQADVA 468
476 LIRVYNPDGTGRTLEAKLHRSQFTVSHHTGAYPLVPPNGHFRFDSWVNOFYSLAPMGTC 535
476 LIRVYNPDGTGRTLEAKLHRSQFTVSHHTGAYPLVPPNGHFRFDSWVNOFYSLAPMGTC 535
469 LIRVYNPDGTGRTLEAKLHRSQFTVSHHTGAYPLVPPNGHFRFDSWVNOFYSLAPMGTC 528
469 LIRVYNPDGTGRTLEAKLHRSQFTVSHHTGAYPLVPPNGHFRFDSWVNOFYSLAPMGTC 528
536 NGRRRRIQ 542
536 NGRRRRIQ 542
529 NGRRRVQ 535
529 NGRRRVQ 535
RESULT 7
Q8V784 PRELIMINARY; PRT; 535 AA.
AC Q8V784;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF2 protein (Capaid).
OS Norwalk virus, and
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OC NCBI_TaxID=11983, 95340;
[1]
RN SEQUENCE FROM N.A.
RP SPECIES=Norwalk virus; STRAIN=UIGII;
RC Kojima S., Kageyama T., Fukushi S., Hoshino F.B., Shinohara M.,
RA Uchida K., Natori K., Takeda N., Katayama K.;
RT "Genogroup-specific primers for detect Norwalk like virus.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RP SPECIES=Norwalk-like virus; STRAIN=Saitama UI;
RC Katayama K.;
```

Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RL SPECIES=Norwalk-like virus; STRAIN=Saitama UI;
RN MEDLINE=22192455; PubMed=12202225;
RC Katayama K., Shirato-Horikoshi H., Kojima S., Kageyama T., Oka T.,
RA Hoshino F.B., Fukushi S., Shinohara M., Uchida K., Suzuki Y.,
RA Gotohori T., Takeda N.;
RT "Phylogenetic Analysis of the Complete Genome of 18 Norwalk-like
RT Viruses."
RL Virology 299:225-239 (2002).
DR EMBL: AB067536; BAB84137.1; -
DR EMBL: AB039775; BAC11816.1; -
DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR008975; Viral_Cap_coat.
DR Pfam: PF00915; Calici_coat.1.
SQ SEQUENCE 535 AA; 58615 MW; 680B2D3162436787 CRC64;

Query Match 70.9%; Score 2028.5; DB 12; Length 535;
Best Local Similarity 71.1%; Pred. No. 3.7e-151;
Matches 389; Conservative 66; Mismatches 75; Indels 17; Gaps 7;
QY 1 MKMASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGOTNIIDPWIRTNFVQAP 60
DB 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
QY 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKIL 120
DB 61 NGEFTVSPRNSPGEVLLNLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKLV 120
QY 121 FAATPPNPLVDNISPAQITMLPHLIVDRTLEPIMTLPDVRNVFHFNNQPPMRMLVA 180
DB 121 FAATPPNPLVDNISPAQITMLPHLIVDRTLEPIMTLPDVRNVFHFNNQPPMRMLVA 180
QY 181 MLYTPLRNSGDDVFTVSCRVLTPDPEFFIYLVPPSVESKTKPFTLPILTISELTNS 240
DB 181 MLYTPLRNSGDDVFTVSCRVLTPDPEFFIYLVPPSVESKTKPFTLPILTISELTNS 240
QY 241 RPPPIEQLYTPNETNVVQNGRCTLGDLGQTTQLLSSAVCFLOGR---TVADNGDN 297
DB 241 RPPVIDELYTPNESLVVQNGRCALDGLRGTTQLLPTAICFRGRINQKVS GENHV 300
QY 298 WDQNLQLYTPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQ-DNVNVSTGEAKNAKGIYI 356
DB 301 WN---MQVTNINGTGFDETPDVPALGTDFSGKLVGVLSQRDHDN-----ACRSHDAVI 352
QY 357 STTSCKFTPKIGSIGLHSITE-HVHPNQSRFTPVGVADENTPQQWVLPHYAGSLAIN 415
DB 353 ATNSAKFTPKLGAIQIGTWEEEDVHINQFTKTPVGLF--ENEGFNQWTLPNYSALTIN 410
QY 416 TNLAPAVAPTFPGEQLLFFRSRVPVCVQGLQGDADFIDCLLPQEWVNFHYQAAAPQADVA 475
DB 411 MGLAPVAPTFPGEKILFRSHIPLKGV--ADPVIDCLLQEWVNFHYQAAAPQADVA 468
QY 476 LIRYVNPDTGRTLFEAKLHRSGLFTVSHGTAYPLVPPNGHFRFDSWVNFYSLAPMGITG 535
DB 469 LIRFTNPDTGRTLFEAKLHRSGLFTVANTGRPIVVPANGYFRFDSWVNFYSLAPMGITG 528
QY 536 NGRRIQ 542
DB 529 NGRREVQ 535

RESULT 8

Q68104 ID Q68104 PRELIMINARY; PRT; 535 AA.
AC Q68104;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Hawaii calicivirus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC Norovirus.
OX NCBI_TaxID=33750;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=Hu/NLV/Hawaii virus/1971/US;
RX MEDLINE=94358485; PubMed=8077710;
RA Lew J.F., Kapikian A.Z., Valdeusio J., Green K.Y.;
RT "Molecular characterization of Hawaii virus and other Norwalk-like
RT viruses: evidence for genetic polymorphism among human
RT caliciviruses."
RL J. Infect. Dis. 170:535-542 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;
RX Lew J.F.;
RT Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;
RX Green K.Y., Sosnovtseva S.;
RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;
RX Pietneva M.A., Sosnovtseva S., Green K.Y.;
RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: U07611; AAB97768.2; -
DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR008975; Viral_Cap_coat.
DR Pfam: PF00915; Calici_coat.1.
SQ SEQUENCE 535 AA; 58741 MW; D0CA252A21084599 CRC64;

Query Match 70.6%; Score 2020.5; DB 12; Length 535;
Best Local Similarity 71.9%; Pred. No. 1.6e-150;
Matches 391; Conservative 59; Mismatches 83; Indels 11; Gaps 6;
QY 1 MKMASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGOTNIIDPWIRTNFVQAP 60
DB 1 MKMASNDAAPSNDGAAGLVPEVNNETMALEPVAGASIAAPLTGQNNIIDPWIRNFVQAP 60
QY 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKIL 120
DB 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKLV 120
QY 121 FAATPPNPLVDNISPAQITMLPHLIVDRTLEPIMTLPDVRNVFHFNNQPPMRMLVA 180
DB 121 FAATPPNPLVDNISPAQITMLPHLIVDRTLEPIMTLPDVRNVFHFNNQPPMRMLVA 180
QY 181 MLYTPLRNSGDDVFTVSCRVLTPDPEFFIYLVPPSVESKTKPFTLPILTISELTNS 240
DB 181 MLYTPLRNSGDDVFTVSCRVLTPDPEFFIYLVPPSVESKTKPFTLPILTISELTNS 240
QY 241 RPPPIEQLYTPNETNVVQNGRCTLGDLGQTTQLLSSAVCFLOGR---TVADNGDN 300
DB 241 RPPVIDELYTPNESLVVQNGRCTLGDLGQTTQLLSSAVCFLOGR---TVADNGDN 300
QY 301 NLLQLYTPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQ-DNVNVSTGEAKNAKGIYISTTS 360
DB 301 NLLQLYTPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQ-DNVNVSTGEAKNAKGIYISTTS 360
QY 361 GKFTPKIGSIGLHSITE-HVHPNQSRFTPVGVADENTP-FQWVLPHYAGSLANTNL 418
DB 357 PKFTPKIGSIGLHSITE-HVHPNQSRFTPVGVADENTP-FQWVLPHYAGSLANTNL 413
QY 419 APVAPVTFPGEQLLFFRSRVPVCVQGLQGDADFIDCLLPQEWVNFHYQAAAPQADVALIR 478
DB 414 APSVPLFPGEQLLFFRSRVPVCVQGLQGDADFIDCLLPQEWVNFHYQAAAPQADVALIR 471
QY 479 YVNPDTGRTLFEAKLHRSGLFTVSHGTAYPLVPPNGHFRFDSWVNFYSLAPMGITGNGR 538
DB 472 YVNPDTGRTLFEAKLHRSGLFTVSHGTAYPLVPPNGHFRFDSWVNFYSLAPMGITGNGR 531
QY 539 RRIQ 542


```
QY 420 PAVAPTFEGEQLLFFRSRVPVCGVQGGQDAFIDCLLPQEWVNHFYQEAAPSQADVALIRY 479
DB 415 PSVAPLFFEGEQLLFFRSRVPVCGGT--SNGAIDCLLPQEWVNHFYQESAPSSTDVALIRY 472
QY 480 VNPDGRTLFEAKLHRSRGPITVSHGTGAYPLVPPNGHFRFDSWVNHQFVSLAPMGNGRR 539
DB 473 TNPDGRLVFEAKLHRSRGPITVSHGTGAYPLVPPNGHFRFDSWVNHQFVSLAPMGNGRR 532
QY 540 RIQ 542
DB 533 RVQ 535

RESULT 11
Q917Y3
ID Q917Y3 PRELIMINARY; PRT; 535 AA.
AC Q917Y3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Westover/302/1994/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171850;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Westover/302/1994/US;
RX MEDLINE=97193806; PubMed=9401391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Westover/302/1994/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
characterized small round-structured viruses involved in outbreaks of
nonbacterial acute gastroenteritis in the United States, 1990 to
1995.";
RL J. Med. Virol. 53:372-383(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Westover/302/1994/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Westover/302/1994/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414418; RAL13007.1;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 535 AA; 58684 MW; 62620468ABB44EC0 CRC64;

Query Match 70.2%; Score 2010.5; DB 12; Length 535;
Best Local Similarity 71.6%; Pred. No. 9.6e-150;
Matches 389; Conservative 61; Mismatches 84; Indels 9; Gaps 6;

QY 1 MKNASNDAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQTNIDPWRITNFVQAP 60
DB 1 MKNASNDAPSNDGAAGLVPEVNNETMALEPVAGASIAAPLTGQNNVIDPWRITNFVQAP 60
QY 61 NGEFTVSPRNSPGEILLNLELPGDLPNLYLAHLSRMYNGYAGGVEVQVLLAGNAFTAGKIL 120
```

```
DB 61 NGEFTVSPRNSPGEILLNLELPGDLPNLYLAHLSRMYNGYAGGVEVQVLLAGNAFTAGKIL 120
QY 121 FAALPPNPLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFVFNNOQPQMRMLVA 180
DB 121 FAALPPHPPIENLSFGQITMFPVHVIDVRLPVLPLPDVRNFFHYNOBEPMRMLVA 180
QY 181 MLYTPLRNSGGDDVFTVSCRVLTPTDPTDFEYILVPPSVESKTPFTLPIITISBELTNS 240
DB 181 MLYTPLRNSGGDDVFTVSCRVLTPTDPTDFEYILVPPSVESKTPFTLPIITISBELTNS 240
QY 241 RFPPIEOLYTPANETNVVQCGRCITLDGELQGTQTLTLLSSAVCFLOGRTVADNGDNMDO 300
DB 241 RFPAPIDELVTSNPEGLVVQNGRSTLDGELGTTLVPPSNICSLGRINAHLSNDQHR 300
QY 301 NLLQLYTPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGIYISTTS 360
DB 301 WNMQVTNANGTPDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGIYISTTS 360
QY 361 GKFTPKIGSIGHLSITEH-VHPNQOSRTPVGVAVDENTPFOQWVLPVHYAGSLANTNLA 419
DB 357 PKFTPKLGSVVLGTWEDRDFDINOPTFTPVGLY--DTDHFNQWALPNYSGALTLMNLA 414
QY 420 PAVAPTFEGEQLLFFRSRVPVCGVQGGQDAFIDCLLPQEWVNHFYQEAAPSQADVALIRY 479
DB 415 PSVAPLFFEGEQLLFFRSRVPVCGGT--SNGAIDCLLPQEWVNHFYQESAPSSTDVALIRY 472
QY 480 VNPDGRTLFEAKLHRSRGPITVSHGTGAYPLVPPNGHFRFDSWVNHQFVSLAPMGNGRR 539
DB 473 TNPDGRLVFEAKLHRSRGPITVSHGTGAYPLVPPNGHFRFDSWVNHQFVSLAPMGNGRR 532
QY 540 RIQ 542
DB 533 RVQ 535

RESULT 12
Q917X4
ID Q917X4 PRELIMINARY; PRT; 535 AA.
AC Q917X4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Port Canaveral/301/1994/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171853;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Port Canaveral/301/1994/US;
RX MEDLINE=97193806; PubMed=9401391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Port Canaveral/301/1994/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
characterized small round-structured viruses involved in outbreaks of
nonbacterial acute gastroenteritis in the United States, 1990 to
1995.";
RL J. Med. Virol. 53:372-383(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Port Canaveral/301/1994/US;
RX MEDLINE=20266071; PubMed=10804147;
```


[illegible]

RA	Ando T., Noel J.S., Fankhauser R.L.; "Genetic classification of 'Norwalk'-like viruses.";
RL	J. Infect. Dis. 181:S336-S348 (2000).
RN	[4]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Hu/NLV/Port Canaveral/301/1994/US;
RA	Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL	Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; AF414421; AAL13016.1; -
DR	InterPro; IPR004005; Calici_coat.
DR	InterPro; IPR008975; Viral_Cap_coat.
DR	Pfam; PF00915; Calici_coat; 1.
SQ	SEQUENCE 535 AA; 58738 MW; 7C29B02DFDC724C6 CRC64;
Query Match 70.1%; Score 2007.5; DB 12; Length 535;	
Best Local Similarity 71.5%; Pred. No. 1.7e-149;	
Matches 388; Conservative 62; Mismatches 84; Indels 9; Gaps 6;	
Qy	1 MKMASNDAAPSDCAASIVPGEINETMPLEPVAGASIAAPVAGQTNIIDPWIRTNFVQAP 60
Dd	1 MKMASNDAAPSDCAAGALVPVENNTWMALEPVAGASIAAPLTGQNVIDPWIRNFVQAP 60
Qy	61 NGFTVSPRNSPGBILLNLGLPDLPNPYLAHLSRMINGYAGGVQVLVLAGNAFTAGKIL 120
Dd	61 NGFTVSPRNSPGBILLNLGLPELNPPFLAHLSRMNGYAGGVQVLLAGNAFTAGKL 120
Qy	121 FAALTPPLVDMSPAQITMLPHLIVDVRTLEPIPTPLDPVRNVFYHENNOQPQRMLVA 180
Dd	121 FAALTPHPPIENLSPQITMFPHVIIDVRTLEPVLPLPDVRNNPFHHNQSPPEPRMLVA 180
Qy	181 MLYTPLRNSGGDDVFTVSCRVLTRPTPDFEFIYLVPVSVESKTKPFTLPILTISLTNS 240
Dd	181 MLYTPLRNSGGDDVFTVSCRVLTRPSPDFDNVLPVPTVESKTKPFTLPILTIGELNS 240
Qy	241 RPPIEOLYTAPNETNVVQCNGRCTLDGELQGTTQLLSSAVCFLOQRTVADNCDWDQ 300
Dd	241 RFPAFIDELYTSNEGVLVVQPONGRPTLDGELLGTTLQVLPNSICSLRGRIHAHSDNQHR 300
Qy	301 NLLQUTYPNGASYDDTDEVPAPLQTDQSGMLYGVLTDQNVNSTGEAKAGIVISTTS 360
Dd	301 WNMQVTNANGTFPDTEDAPAPLGTPDFLANIYGVTSQRNPD-NTCRAHD--GI-LATWS 356
Qy	361 GKFTPKIGSIGLHSITEH-VHPNQOSRFTPVGVADVDENTFFQQWLPHYGSIALNTNLA 419
Dd	357 PKFTPLGSGVLGTWEDRFDINQTRTFPVGLY--DTDHFNQVALPNYSICALTLNMILA 414
Qy	420 PAVAFTFGQEQLLFRRSRVPCVQGLQGDAFIDCLLPQEWANHYEQEAAPSQADVALIRY 479
Dd	415 PSVAPLPFCEQLLFFRSHTPLKGKT--SNGAIDCLLPQEWQHIFYQESAPSTDVALIRY 472
Qy	480 VNPDTGRITFEAKLHRSGITVTSHTGAYPLVPPNGHFRFDSVMNGFYSLAPMGTCNGRR 539
Dd	473 TMPDTRVLFEAKLHQRFQFITVANSGSRPLVPPNGGYFRFDSVMNGFYSLAPMGTCNGRR 532
Qy	540 RIQ 542
Dd	533 RVQ 535
RESULT 13	
ID	Q916E4 PRELIMINARY; PRT; 535 AA.
AC	Q916E4;
DT	01-DEC-2001 (TEMBLrel. 19, Created)
DT	01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT	01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE	Capsid protein.
OS	Human calicivirus NLV/Wiesbaden 294/01/DE.
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC	Norovirus.
OX	NCBI_TaxID=173922;
RP	SEQUENCE FROM N.A. [1]

RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Richmond/283/1994/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
RA Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995.";
RL J. Med. Virol. 53:372-383(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Richmond/283/1994/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk'-like viruses.";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Richmond/283/1994/US;
RA Ando T., Seto Y., Noel J.S., Glass R.I., Fankhauser R.L.;
RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF414419; AAL13010.1; -;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 535 AA; 58742 MW; 4EDACABC8B22F80 CRC64;

Query Match 70.0%; Score 2003.5; DB 12; Length 535;
Best Local Similarity 71.5%; Pred. No. 3.4e-149; Indels 9; Gaps 6;
Matches 388; Conservative 62; Mismatches 84;
QY 1 MKMASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQNTIIDPWIRTFVQAP 60
DB 1 MKMASNDAAPSNDGAAGLVPEVNNETMALEPVAGASIAAPLTGQNNVIDPWIRMFVQAP 60
QY 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKIL 120
DB 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKIL 120
QY 121 FAALPPNPLVDMSIPAQITMLPHLIVDVRTLEPIMTLPDVRNVFHHNQOPQPRMLVA 180
DB 121 FAALPPHPPIENLSPQITMFPVHVIDVRLTLEPVLPLPDVRNNFFHYNQPEPRMLVA 180
QY 181 MLYTPLRSNGSGDDVFTVSCRVLTRPTDPFEIYLVPPSVESKTKPFTLPILTIGELTNS 240
DB 181 MLYTPLRSNGSGDDVFTVSCRVLTRPTDPFEIYLVPPSVESKTKPFTLPILTIGELTNS 240
QY 241 RFPPIEQLYTAPNETNVVQCGRCTLDELGGTTLQSSAVCFLOQRTVADNGDNWDQ 300
DB 241 RFPPIEQLYTAPNETNVVQCGRCTLDELGGTTLQSSAVCFLOQRTVADNGDNWDQ 300
QY 301 NLLQLYTPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAGIYISTTS 360
DB 301 NLLQLYTPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAGIYISTTS 360
QY 361 KFTPKIGSLGHSITEH-VHPNQSRPTPGVAVDENTPFOQWVLPYHAGSLALNTNLA 419
DB 361 KFTPKIGSLGHSITEH-VHPNQSRPTPGVAVDENTPFOQWVLPYHAGSLALNTNLA 419
QY 420 PAVAPTFPGEQLLFFRSRVPVCQGLQGDADFIDCLLPQEWNVHFOEAPSOADVALIRY 479
DB 420 PAVAPTFPGEQLLFFRSRVPVCQGLQGDADFIDCLLPQEWNVHFOEAPSOADVALIRY 479
QY 479 VNPDTGRTLFEAKLHRSRGTITVSHTGAYPLVVPNGHFRFDSWNQFYSLAPMGTCNGRR 539
DB 479 VNPDTGRTLFEAKLHRSRGTITVSHTGAYPLVVPNGHFRFDSWNQFYSLAPMGTCNGRR 539

QY 540 RIQ 542
DB 533 RVQ 535

RESULT 15

QY16E6 PRELIMINARY; PRT; 535 AA.
AC Q916E6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/Dillingen 391/01/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=173920;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/Dillingen 391/01/DE;
RA Kuenkel U., Hoehne M., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated
RT with Norwalk-like viruses in Germany.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF425767; AAL18860.1; -;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 535 AA; 58685 MW; 0520E521A786734A CRC64;

Query Match 69.9%; Score 2002.5; DB 12; Length 535;
Best Local Similarity 71.3%; Pred. No. 4.1e-149;
Matches 387; Conservative 62; Mismatches 85; Indels 9; Gaps 6;

QY 1 MKMASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQNTIIDPWIRTFVQAP 60
DB 1 MKMASNDAAPSNDGAAGLVPEVNNETMALEPVAGASIAAPLTGQNNVIDPWIRMFVQAP 60
QY 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKIL 120
DB 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKIL 120
QY 121 FAALPPNPLVDMSIPAQITMLPHLIVDVRTLEPIMTLPDVRNVFHHNQOPQPRMLVA 180
DB 121 FAALPPHPPIENLSPQITMFPVHVIDVRLTLEPVLPLPDVRNNFFHYNQPEPRMLVA 180
QY 181 MLYTPLRSNGSGDDVFTVSCRVLTRPTDPFEIYLVPPSVESKTKPFTLPILTIGELTNS 240
DB 181 MLYTPLRSNGSGDDVFTVSCRVLTRPTDPFEIYLVPPSVESKTKPFTLPILTIGELTNS 240
QY 241 RFPPIEQLYTAPNETNVVQCGRCTLDELGGTTLQSSAVCFLOQRTVADNGDNWDQ 300
DB 241 RFPPIEQLYTAPNETNVVQCGRCTLDELGGTTLQSSAVCFLOQRTVADNGDNWDQ 300
QY 301 NLLQLYTPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAGIYISTTS 360
DB 301 NLLQLYTPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAGIYISTTS 360
QY 361 KFTPKIGSLGHSITEH-VHPNQSRPTPGVAVDENTPFOQWVLPYHAGSLALNTNLA 419
DB 361 KFTPKIGSLGHSITEH-VHPNQSRPTPGVAVDENTPFOQWVLPYHAGSLALNTNLA 419
QY 420 PAVAPTFPGEQLLFFRSRVPVCQGLQGDADFIDCLLPQEWNVHFOEAPSOADVALIRY 479
DB 420 PAVAPTFPGEQLLFFRSRVPVCQGLQGDADFIDCLLPQEWNVHFOEAPSOADVALIRY 479
QY 479 VNPDTGRTLFEAKLHRSRGTITVSHTGAYPLVVPNGHFRFDSWNQFYSLAPMGTCNGRR 539
DB 479 VNPDTGRTLFEAKLHRSRGTITVSHTGAYPLVVPNGHFRFDSWNQFYSLAPMGTCNGRR 539

QY 540 RIQ 542
DB 533 RVQ 535

Wed Jun 2 09:13:35 2004

Db 533 RVQ 535

Search completed: June 1, 2004, 13:53:21
Job time : 32.556 secs

us-09-926-799-9.rsp

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 47.2483 Seconds
(without alignments)
3289.030 Million cell updates/sec

Title: US-09-926-799-10

Perfect score: 2906

Sequence: 1 MKMASNDAPNSDGAANLVP.....GNQPYTLAPMGSGGRRRAQ 550

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2906	100.0	550	4	AAB49709 Small rou
2	2124	73.1	548	4	AAB49705 Small rou
3	2074	71.4	548	5	AAU91272 Norwalk v
4	2062	71.0	535	4	AAB49707 Small rou
5	2001	68.9	540	4	AAB49706 Small rou
6	1872	64.4	542	4	AAB49708 Small rou
7	1834.5	63.1	539	4	AAB49704 Small rou
8	1613.5	55.5	541	4	AAB49710 Small rou
9	1154.5	39.7	545	4	AAB49700 Small rou
10	1141	39.3	530	2	AAR50972 Norwalk v
11	1141	39.3	530	7	ADC72176 Norwalk v
12	1131	38.9	530	4	AAB49701 Small rou
13	1124.5	38.7	544	4	AAB49703 Small rou
14	1097	37.7	530	2	AAR57091 Small rou
15	1090.5	37.5	546	4	AAB49702 Small rou
16	285	9.8	579	2	AAW08143 RHDF caps
17	280.5	9.7	547	4	AAW50108 Feline ca
18	280.5	9.7	668	4	AAB67462 Amino aci
19	280.5	9.7	671	4	AAW50107 Feline ca
20	276.5	9.5	622	4	AAB47045 Feline Ca
21	276.5	9.5	623	4	AAB47044 Feline Ca
22	276.5	9.5	623	4	AAB47043 Feline Ca
23	272.5	9.4	669	4	AAB67461 Amino aci
24	270.5	9.3	668	2	AAR10686 Feline ca
25	270.5	9.3	668	4	AAE04304 Feline ca

26	179.5	6.2	40	5	AAU91274	AAU91274 Norwalk v
27	168	5.8	40	5	AAU91273	AAU91273 Norwalk v
28	138.5	4.4	934	1	AAU20016	AAU20016 Sequence
29	124.5	4.3	1147	5	ABB76724	ABB76724 Foot and
30	121	4.2	1344	4	AAU91275	AAU91275 Norwalk v
31	120	4.1	2608	5	ABG70209	ABG70209 Human pre
32	119.5	4.1	1036	6	ABU70541	ABU70541 Human adi
33	119	4.1	2647	2	AAW19349	AAW19349 Human fil
34	119	4.1	2647	2	AAW19349	AAW19349 Human fil
35	119	4.1	2647	2	AAW19349	AAW19349 Human fil
36	119	4.1	2647	2	AAW19349	AAW19349 Human fil
37	117	4.0	4618	4	AAW39043	AAW39043 Human pol
38	113	3.9	701	4	ABB57994	ABB57994 Drosophil
39	112.5	3.9	568	4	ABB64772	ABB64772 Drosophil
40	109.5	3.8	1079	2	AAU03163	AAU03163 MuLV reve
41	109.5	3.8	1199	3	AAU12994	AAU12994 MuLV rever
42	109.5	3.8	1224	2	AAU17947	AAU17947 MoMLV pol
43	109.5	3.8	1737	3	AAU10044	AAU10044 MoMLV 989-
44	107.5	3.7	1958	4	ABG21921	ABG21921 Novel hum
45	107	3.7	557	5	ABP61060	ABP61060 Lactobaci

ALIGNMENTS

RESULT 1
AAB49709
ID AAB49709 standard; protein; 550 AA.
XX
AC AAB49709;
XX
DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 10.
XX
KW Small round structured virus; SRSV; food poisoning.
XX
OS Small round structured virus.
XX
PN WO200079280-A1.
XX
FD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000WO-JP004095.
XX
PR 22-JUN-1999; 99JP-00175928.
XX
PA (NTINA-) JAPAN NAT INST INFECTIOUS DISEASES.
FA (DENK-) DENKA SEIKEN KK.
XX
Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
WPI; 2001-080848/09.
N-PSDB; AAF29150.
XX
Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.
XX
Claim 1; Page 62-64; 84pp; Japanese.
XX
This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks
XX
SQ Sequence 550 AA;
Query Match 100.0%; Score 2906; DB 4; Length 550;
Best Local Similarity 100.0%; Pred. No. 1.4e-260;

PT	Strain	Match	Conservative	Mismatches	Indels	Gaps	Length
PT	Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.	550	0	0	0	0	548
XX	Claim 1; Page 52-54; 84pp; Japanese.						
PS	This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks						
XX	Sequence 548 AA;						
SQ	Query Match 73.1%; Score 2124; DB 4; Length 548; Best Local Similarity 71.4%; Pred. No. 6.3e-186; Indels 12; Gaps 4; Matches 396; Conservative 72; Mismatches 75;						
Qy	1 MKMASNDAAAPSDNGAANLVPEANDEVMALPEVVGASIAAPVVGQNIIDPWIRENFVQAP 60						
Db	1 MKMASNDAAAPSDNGAANLVPEANDEVMALPEVVGASIAAPVVGQNIIDPWIRENFVQAP 60						
Qy	61 QGEFTVSPRNSPGEMLNLELPELVLSHLSRMVNGYAGGMQVQVVLGNAFTAGKII 120						
Db	61 QGEFTVSPRNSPGEMLNLELPELVLSHLSRMVNGYAGGMQVQVVLGNAFTAGKII 120						
Qy	121 FAAPPHPPVENISAAQITMCPHIVDVRLPEVLLPLPDIRNPFHYNOENTPRMLVA 180						
Db	121 FAAPPHPPVENISAAQITMCPHIVDVRLPEVLLPLPDIRNPFHYNOENTPRMLVA 180						
Qy	181 MLYTPLRANSGEDVFTVSCRVLTRPADFEFTLVPTVESKTKPFTLPIITLGEISNS 239						
Db	181 MLYTPLRANSGEDVFTVSCRVLTRPADFEFTLVPTVESKTKPFTLPIITLGEISNS 239						
Qy	240 RFPAAIDMLYTDNESHIVQPNQGRCTLDGTTQTLVPTQICAPRGTLISQTAARAS 299						
Db	240 RFPAAIDMLYTDNESHIVQPNQGRCTLDGTTQTLVPTQICAPRGTLISQTAARAS 299						
Qy	300 TD-SQPQARNHPLHVQVKNLDGTQYDPTDDIPAVLGAIDFKGTVFGVASORDVSGQOEQ 358						
Db	300 TD-SQPQARNHPLHVQVKNLDGTQYDPTDDIPAVLGAIDFKGTVFGVASORDVSGQOEQ 358						
Qy	359 HYATRAHEAHIDTDPKYAPKGLTILKSGDDPNTNQPIRFTPVGMGNWQWEL 415						
Db	359 HYATRAHEAHIDTDPKYAPKGLTILKSGDDPNTNQPIRFTPVGMGNWQWEL 415						
Qy	416 PDYSGQFTHNNLAPAVAPNPFGEQLLFFRSQPSGSGRSNGVLDCLVPOEWQHFYQES 473						
Db	416 PDYSGQFTHNNLAPAVAPNPFGEQLLFFRSQPSGSGRSNGVLDCLVPOEWQHFYQES 473						
Qy	474 APOQTOVALVRYVNPDTGRNIFEAHLHREGFLTVANCNNPIVVPNGYFEEAWGNQFY 535						
Db	474 APOQTOVALVRYVNPDTGRNIFEAHLHREGFLTVANCNNPIVVPNGYFEEAWGNQFY 535						
Qy	536 TLAPMGSGQGRRAQ 550						
Db	536 TLAPMGSGQGRRAQ 550						
Qy	541 GSGQGRRAQ 550						
Db	541 GSGQGRRAQ 550						
RESULT 2							
AAAB49705							
ID	AAAB49705 standard; protein; 548 AA.						
XX	AAAB49705;						
XX	04-APR-2001 (first entry)						
DE	Small round structured virus protein SEQ ID 6.						
XX	Small round structured virus; SRSV; food poisoning.						
XX	Small round structured virus.						
XX	WO200079280-A1.						
XX	28-DEC-2000.						
XX	22-JUN-2000; 2000WO-JP004095.						
XX	22-JUN-1999; 99JP-00175928.						
XX	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.						
XX	(DENK-) DENKA SEIKEN KK.						
XX	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;						
XX	WPI; 2001-080848/09.						
XX	N-PSDB; AAF29146.						


```
PS Claim 1; Page 59-61; 84pp; Japanese.
XX
CC This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX
SQ Sequence 542 AA;

Query Match 64.4%; Score 1872; DB 4; Length 542;
Best Local Similarity 64.9%; Pred. No. 1.6e-164;
Matches 361; Conservative 74; Mismatches 101; Indels 20; Gaps 7;

Qy 1 MKMASNDAAPSNDGAANLVPEANDSWALEPVGASIAAPVVGQNIIDPWIRENFVQAP 60
Db 1 MKMASNDAAPSNDGAANLVPEANDSWALEPVGASIAAPVVGQNIIDPWIRENFVQAP 60

Qy 61 QGEFTVSPRNSPGEMLLNLLELGPENLPYLSHLSRMVYAGGMOVQVVLGNAFTAGKII 120
Db 61 NGEFTVSPRNSPGEMLLNLLELGPENLPYLSHLSRMVYAGGVEQVLLAGNAFTAGKIL 120

Qy 121 FAAVPPHPVENISAAQITMCQPHVIVDVRQLEPVLPLDPIRNFPHYQNTPRMLVA 180
Db 121 FAAVPPHPVENISAAQITMCQPHVIVDVRQLEPVLPLDPIRNFPHYQNTPRMLVA 180

Qy 121 FAAVPPHPVENISAAQITMCQPHVIVDVRQLEPVLPLDPIRNFPHYQNTPRMLVA 180
Db 121 FAAVPPHPVENISAAQITMCQPHVIVDVRQLEPVLPLDPIRNFPHYQNTPRMLVA 180

Qy 181 MLYTPLRAN-SEGEDVFTVSCRVLTRPADPDEFTFLVPPTVESKTKPFTLPILTILGELS 239
Db 181 MLYTPLRAN-SEGEDVFTVSCRVLTRPADPDEFTFLVPPTVESKTKPFTLPILTILGELS 239

Qy 240 RFPAAIDMLYTDPNESIVVQPNQGRCTLDGTLOQTTLQVPTQICAFRGTLSISQARAADS 299
Db 240 RFPAAIDMLYTDPNESIVVQPNQGRCTLDGTLOQTTLQVPTQICAFRGTLSISQARAADS 299

Qy 241 RFPPIELKLYTGSSAFVQPNQGRCTLDGTLOQTTLQVPTQICAFRGTLSISQARAADS 299
Db 241 RFPPIELKLYTGSSAFVQPNQGRCTLDGTLOQTTLQVPTQICAFRGTLSISQARAADS 299

Qy 300 TDSQARARNHPLHVQVKNLDGTQYDPTDDIPAVLGAIDFKGTVFGVASQRDVSGQOEGH 359
Db 300 TDSQARARNHPLHVQVKNLDGTQYDPTDDIPAVLGAIDFKGTVFGVASQRDVSGQOEGH 359

Qy 298 WDQNLQLTYP-----NGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGE-- 347
Db 298 WDQNLQLTYP-----NGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGE-- 347

Qy 360 YATRAHEAHITDTPKYAPKLGTLIKSGSDDFNTNQPITRTPVCMG-DNN--NRWELP 416
Db 360 YATRAHEAHITDTPKYAPKLGTLIKSGSDDFNTNQPITRTPVCMG-DNN--NRWELP 416

Qy 348 -AKNAKGIYITTSCKTPKIGSLHSITHEVHPNQSRFTFPGVAVDENTPFQOWVLP 406
Db 348 -AKNAKGIYITTSCKTPKIGSLHSITHEVHPNQSRFTFPGVAVDENTPFQOWVLP 406

Qy 417 DYSGLTLNMLNLAAPSFPGERILFPRISIVPSAGGY--GSGYIDCLIPQEWVQHFQYE 474
Db 417 DYSGLTLNMLNLAAPSFPGERILFPRISIVPSAGGY--GSGYIDCLIPQEWVQHFQYE 474

Qy 407 HYAGSLALNTNLAPAVPTFGEQLLFFRSRVPVQVQLOGQDAFIDCLLPQEWVQHFQ 466
Db 407 HYAGSLALNTNLAPAVPTFGEQLLFFRSRVPVQVQLOGQDAFIDCLLPQEWVQHFQ 466

Qy 475 AAPQSQAVLVRYNPDTRGRNIFEAKLHREGFLTVANGCNPNIVVPPNGYFRFPAWGN 534
Db 475 AAPQSQAVLVRYNPDTRGRNIFEAKLHREGFLTVANGCNPNIVVPPNGYFRFPAWGN 534

Qy 467 AAPSQADVALLIRYVNPDTGRVLFCKLHRSKGYITVSHGTGAYPLVVPNGHFRFDSWV 526
Db 467 AAPSQADVALLIRYVNPDTGRVLFCKLHRSKGYITVSHGTGAYPLVVPNGHFRFDSWV 526

Qy 535 YTLAPMGSGQGRRAA 550
Db 535 YTLAPMGSGQGRRAA 550

Qy 527 YSLAPMGSGNGRRRIQ 542
Db 527 YSLAPMGSGNGRRRIQ 542

RESULT 7
AAB49704
ID AAB49704 standard; protein; 539 AA.
XX
AC AAB49704;
XX
DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 5.
XX
KW Small round structured virus; SRSV; food poisoning.
XX
OS Small round structured virus.
XX
FN WO200079280-A1.
XX
```

```
PD 28-DEC-2000.
XX
XX 22-JUN-2000; 2000WO-JP004095.
XX
XX 22-JUN-1999; 99JP-00175928.
XX
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX
XX (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI; 2001-080848/09.
XX
XX N-FSDB; AAF29145.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX strains for investigation of food poisoning outbreaks, contains
XX antibodies.
XX
XX Claim 1; Page 50-52; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 539 AA;

Query Match 63.1%; Score 1834.5; DB 4; Length 539;
Best Local Similarity 62.5%; Pred. No. 4.9e-161;
Matches 348; Conservative 85; Mismatches 97; Indels 27; Gaps 7;

Qy 1 MKMASNDAAPSNDGAANLVPEANDSWALEPVGASIAAPVVGQNIIDPWIRENFVQAP 60
Db 1 MKMASNDAAPSNDGAANLVPEANDSWALEPVGASIAAPVVGQNIIDPWIRENFVQAP 60

Qy 61 QGEFTVSPRNSPGEMLLNLLELGPENLPYLSHLSRMVYAGGMOVQVVLGNAFTAGKII 120
Db 61 NGEFTVSPRNSPGEMLLNLLELGPENLPYLSHLSRMVYAGGVEQVLLAGNAFTAGKII 120

Qy 121 FAAVPPHPVENISAAQITMCQPHVIVDVRQLEPVLPLDPIRNFPHYQNTPRMLVA 180
Db 121 FAAVPPHPVENISAAQITMCQPHVIVDVRQLEPVLPLDPIRNFPHYQNTPRMLVA 180

Qy 181 MLYTPLRAN-SEGEDVFTVSCRVLTRPADPDEFTFLVPPTVESKTKPFTLPILTILGELS 239
Db 181 MLYTPLRAN-SEGEDVFTVSCRVLTRPADPDEFTFLVPPTVESKTKPFTLPILTILGELS 239

Qy 240 RFPAAIDMLYTDPNESIVVQPNQGRCTLDGTLOQTTLQVPTQICAFRGTLSISQARAADS 299
Db 240 RFPAAIDMLYTDPNESIVVQPNQGRCTLDGTLOQTTLQVPTQICAFRGTLSISQARAADS 299

Qy 241 RFPPIELKLYTGSSAFVQPNQGRCTLDGTLOQTTLQVPTQICAFRGTLSISQARAADS 299
Db 241 RFPPIELKLYTGSSAFVQPNQGRCTLDGTLOQTTLQVPTQICAFRGTLSISQARAADS 299

Qy 300 TDSQARARNHPLHVQVKNLDGTQYDPTDDIPAVLGAIDFKGTVFGVASQRDVSGQOEGH 359
Db 300 TDSQARARNHPLHVQVKNLDGTQYDPTDDIPAVLGAIDFKGTVFGVASQRDVSGQOEGH 359

Qy 289 -DVTHIAGSHDYTMNLASQWNSNDPTEEPAPLGTQDFVGVKIQGLTQT-----TREDG- 342
Db 289 -DVTHIAGSHDYTMNLASQWNSNDPTEEPAPLGTQDFVGVKIQGLTQT-----TREDG- 342

Qy 360 YATRAHEAHITDTPKYAPKLGTLIKSG-SDDENTNQPITRTPVCMG---GDNW---RQ 412
Db 360 YATRAHEAHITDTPKYAPKLGTLIKSG-SDDENTNQPITRTPVCMG---GDNW---RQ 412

Qy 343 -STRAHKATVSTGSHVFTPKLGSVQYTTDNNDFQGTQNTKFTFVGVQIDGNHNEPQQ 401
Db 343 -STRAHKATVSTGSHVFTPKLGSVQYTTDNNDFQGTQNTKFTFVGVQIDGNHNEPQQ 401

Qy 413 WELPDYSGRLTLNMLNLAAPSFPGERILFPRISIVPSAGGYIDCLIPQEWVQHFY 472
Db 413 WELPDYSGRLTLNMLNLAAPSFPGERILFPRISIVPSAGGYIDCLIPQEWVQHFY 472

Qy 402 WYLENYSGRTHNVHLAPAVPTFGEQLLFFRSRVPVQVQLOGQDAFIDCLLPQEWVQHF 461
Db 402 WYLENYSGRTHNVHLAPAVPTFGEQLLFFRSRVPVQVQLOGQDAFIDCLLPQEWVQHF 461

Qy 473 QEAAPSQAVLVRYNPDTRGRNIFEAKLHREGFLTVANGCNPNIVVPPNGYFRFPAWGN 532
Db 473 QEAAPSQAVLVRYNPDTRGRNIFEAKLHREGFLTVANGCNPNIVVPPNGYFRFPAWGN 532

Qy 462 QEAAPSQADVALLIRYVNPDTGRVLFCKLHRSKGYITVSHGTGAYPLVVPNGHFRFDSWV 521
Db 462 QEAAPSQADVALLIRYVNPDTGRVLFCKLHRSKGYITVSHGTGAYPLVVPNGHFRFDSWV 521

Qy 533 QFYTLAPMGSGQGRRAA 549
Db 533 QFYTLAPMGSGQGRRAA 549

Qy 522 QFYTLAPMGSGNGRRRIQ 538
Db 522 QFYTLAPMGSGNGRRRIQ 538
```

	Db	296	-----DVSRYMDMEITNTDGTPTEDTPGPIGPSDFQGLIFGVASQKNEQP----	347
	Qy	360	YATRAHEAHIDT-----TDPKYAPK---LGTILIKS-----	387
	Dd	348	-ATRAHEAIIINTGGDLHLCPOISSSIYLTSPLNLRCTNPQLPQSGLRGTILIRSDNGHC	406
	Qy	388	----GSDDFNTNQIIRFTVPMGDN---NWROWELPDYSGRLLTNMNLAPAVSPSPFCER	440
	Dd	407	HDMVGTSPTTPTWFQQWRRCRSRSCSGSHRYPPVVMNRVTM-----	450
	Qy	441	ILFRSIVPSAGYGSGYIDCL--IPQ---EWQHIFYOEAAAPSQAVALRVYNPDTGRN	495
	Dd	451	-----IVLS---HKSGFSTSTRKLPLNLRW-----PLIRFINDDTCRV	486
	Qy	496	IFEAKLHREGFLTVANGNNPIVVPPNGYFRFEAMGNQFYTLAPNGSQGRRRAQ	550
	Dd	487	LFEARLHKQGFITVAHTGDNPIVMPNGYFRFEAWNFYSIAPYVGTGKRVRV	541
			RESULT 9	
	ID	AAB49700		
	XX	AAB49700 standard; protein; 545 AA.		
	AC	AAB49700;		
	XX			
	DT	04-APR-2001 (first entry)		
	XX			
	DE	Small round structured virus protein SEQ ID 1.		
	XX			
	KW	Small round structured virus; SRSV; food poisoning.		
	XX			
	OS	Small round structured virus.		
	XX			
	PN	WO200079280-A1.		
	XX			
	PD	28-DEC-2000.		
	XX			
	PF	22-JUN-2000; 2000WO-JP004095.		
	XX			
	PR	22-JUN-1999; 99JP-00175928.		
	XX	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.		
	PA	(DENK-) DENKA SEIKEN KK.		
	XX	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;		
	PI			
	XX	WPI; 2001-080848/09.		
	DR	N-PSDB; AAF29141.		
	XX			
	DR	Kit for the detection and typing of small round-structured virus (SRSV)		
	XX	strains for investigation of food poisoning outbreaks, contains		
	PT	antibodies.		
	XX			
	PT	Claim 1; Page 40-42; 84pp; Japanese.		
	XX	This invention relates to a kit for the detection and typing of small		
	CC	round structured virus (SRSV) strains. The kit contains antibodies		
	XX	directed against peptides represented in sequences AAB49700 - AAB49710,		
	CC	which are each SRSV strain specific. Polynucleotide sequences AAF20141 -		
	XX	AAF20151 represent cDNA encoding the strain specific proteins. The kit is		
	CC	used for detecting and typing strains of SRSV in order to prevent the		
	XX	spread of infection and to examine the epidemiology of outbreaks		
	CC			
	Sequence 541 AA;			
	Query Match	55.5%; Score 1613.5; DB 4; Length 541;		
	Best Local Similarity	55.3%; Pred. No. 1.7e-140;		
	Matches 329; Conservative	77; Mismatches 90; Indels 99; Gaps 12;		
	Qy	1 MKMASNDAAPSNDGAANLVPEANDVEWMALEPVVUGASTAAPVGOQNIIDPWIRENVQAP	60	
	Dd	1 MKMASNDAAPSNDGAAGLVPEINNEVMPLEPVAGASLATVPVGOQNIIDPWIRNNEVQAP	60	
	Qy	61 QGEFTVSRNSPGEMLNLELPNELPYLSHLRMVNYAGGMQVQVVLGNAFTAGKII	120	
	Dd	61 AGEFTVSRNSPGETILLDLELPDNLNPLYLAHARYNHAGGMEVQIVLAGNAFTAGKII	120	
	Qy	121 FFAVPHFPVENISAQITMCIPHVIDVRQLEPVLPLDIDRNRFHYNOENTPRMLVA	180	
	Dd	121 FFAIIPGPFFYENLSFSQITMCIPHVIDVRQLEPVLPLDIDRNRFHYNOENDPKRLVA	180	
	Qy	181 MLTYPLRA-NSEDEVFTVSCRVLTPAPDFEFTFLVPTVESKTPTLTILTGLBLSNS	239	
	Dd	181 MLTYPLRANSNGDDVFTVSCRVLTKPSDFEFTFLVPTVESKTQFALLIKISETWS	240	
	Qy	240 RPPAAIDMLYTDPNESIVVQPNGRCTLDIGTLQGTQTOLVPTQICAFRTGLISQTAARDS	299	
	Dd	241 RFPVPDVMTARNENOVQPNGRVTLDELGLTGTTPLLANVICKFGEVIANKG-----	295	
	Qy	300 TDSFQPARNHLHVQVKNLIDGTQYDPTDDIIPAVLGAIIDFKTVFGVASQRDVSVQEQGH	359	


```
Qy 56 FVQAPQGEFTVSPRNSPCEMLNLLNLELGPENLPYLSHLRMYNGYAGGMQVQVVLGNAFT 115
Db 60 YVQAPQGEFTTSPNNTPGDILFDLQGLPHLPFLSLAQMYNGWGNMKVKVLLAGNAFT 119
Qy 116 AKGIIFAAVPHPPFVENISAAQITMCPIHVIVDVRQLEPVLPLDIRNRPFFHYNQENTPR 175
Db 120 AKGIISICIPGFAAQNISIAQATMFPHVIADVRLEPIEVLBEDVRNVLPH-NNDNAFT 178
Qy 176 MRLVAMLYTPLRA---NSEDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPFTLPLT 232
Db 179 MRLVCMLYTPLRASGSSGTDFFVAGRVLCPSDFSLFLVPPNVQKTKPSVPLP 238
Qy 233 LGEISNSRFPAAIDMLYTDPNESIVVQPNQGRCTLDGTLQGTQLVPTQICAPRGTLLISQ 292
Db 239 LNTLSNSRVPSLRSKMMVSRDHQMVQFQNGRVTLDGQLQGTTPTSASQLCKIRGSVPH- 297
Qy 293 TARAADSTDSQARARNHPLHVQKNLDGTDQDPTDDIPAVLGADTF-----KGTV 342
Db 298 -----ANGNGYNLTDELGSPYHAFES-PAPIGFDPDLGECDMHEASPTTQ 342
Qy 343 FGVASQ-RDVSGQOQOQGHYATRAHEAHIDTDPKYAPKLGTLILIKSGSD-DFNTNOPIRF 400
Db 343 FNTGDVVKQINVKQESA-----FAPHLGTLQADGLSDVSVNTMIAKL 385
Qy 401 ---TPVGMGD-NNRQWELPDYSGRLTLMNMLAPAVSPSPGERILFPRSVPSAGGYGS 456
Db 386 GWSPVSDGRGDVDPWVPIPRYGSTLTEAAQLAPPIYPPGFEAIVPFMSDFPIAHGTNG 445
Qy 457 GYIDCLIPQWVQHFYQEAAPSQSAVALRVYNPDGTGRNIFEAKLHREGFLTVA--NCGN 514
Db 446 LSVPTTIPQEFVTHFVNSQAPTRGEAALLHYLDPDTHRNLEGEFKLYPEGFMTCPVNSSGT 505
Qy 515 NPVIVPPNGVFRFEANGNOFTYTLAPMGS 542
Db 506 GPQLIPNGVFRFVSWRSRFPQLKPVGT 533

RESULT 10
AAR50972
ID AAR50972 standard; protein; 530 AA.
XX
AC AAR50972;
XX
XX 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 05-OCT-1994 (first entry)
XX
XX Norwalk virus strain 8FIIa protein (encoded by ORF2).
XX
XX Norwalk virus; pathogen; acute gastroenteritis; food poisoning;
XX seafood contamination; diagnostic assay; calicivirus; small round virus.
XX
XX Norwalk virus; (strain 8FIIa).
XX
XX WO9405700-A2.
XX
XX 17-MAR-1994.
XX
XX 07-SEP-1993; 93WO-US008447.
XX
XX 07-SEP-1992; 92US-00941365.
XX
XX (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
XX Matson DO, Estes MK, Jiang X, Graham DY;
XX
XX WPI; 1994-101125/12.
XX
XX N-PSDB; AAQ56826.
XX
XX DNA from Norwalk and related viruses - used for preparing prods. for use
XX in diagnostic assays, detection and vaccines for Norwalk and related
XX viruses.
```

```
XX Claim 14; Page 68-70; 156pp; English.
XX
XX The Norwalk virus was isolated from stool samples from adult volunteers
XX infected with safety tested Norwalk virus strain 8FIIa. The coding
XX sequence is useful for the design of probes for use in diagnostic assays
XX for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct
XX PN field.) (Updated on 16-OCT-2003 to standardise OS field)
XX
XX Sequence 530 AA;
XX
XX Query Match 39.3%; Score 1141; DB 2; Length 530;
XX Best Local Similarity 43.4%; Pred. No. 1.3e-96;
XX Matches 249; Conservative 78; Mismatches 169; Indels 78; Gaps 12;
Qy 1 MKMASNDAAAPNDGAA---NLVPEAN-DEVMALEPVVVGASIAAPVVGQONIIDPWIRENF 56
Db 1 MMMAKSKATSDVSDGASGAGQLVPEVNASDPLAMPDVAGSSTAVATAGQVNFIDPWINNF 60
Qy 57 VQAPQGEFTVSPRNSPCEMLNLLNLELGPENLPYLSHLRMYNGYAGGMQVQVVLGNAFTA 116
Db 61 VQAPQGEFTTSPNNTPGDILFDLQGLPHLPFLSLAQMYNGWGNMKVRVRLAGNAFTA 120
Qy 117 GKIIIFAAVPHPPFVENISAAQITMCPIHVIVDVRQLEPVLPLDIRNRPFFHYNQENTPRM 176
Db 121 GKIIIVSCIPPGFGSHNLTIAQTLFPHVIADVRLEPIEVLBEDVRNVLPHNDRNQOTM 180
Qy 177 RLVAMLYTPLRANGSE-DVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPFTLPLTILGE 235
Db 181 RLVCMLYTPLTGTGGTSDSFVAGRVMTCPSPDNFLFLVPTVEQKTRPFTLPLNPLSS 240
Qy 236 LNSRFPAAIDMLYTDPNESIVVQPNQGRCTLDGTLQGTQLVPTQICAPRGTLLISOTAR 295
Db 241 LNSRAFLPISSMGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGT----- 293
Qy 296 AADSTDSQARARNHPLHVQKNLDGTDQDPTDDIPAVLGADTFKG-----TVFGVASQ 348
Db 294 -----SNGTVINLTDELGTPTFPFPEG-PAPIGFDPDLGECDMHINMTQGHSSQ 340
Qy 349 RDVSGQOQOQGHYATRAHEAHIDTDPKYAPKLGTLILIKS-GSDDFNNTNOPIRFTPVGMGD 407
Db 341 T-----QYVDVTTPTDFVPHLGSIQANGISGNY-----VG 372
Qy 408 NNWRQ-----WELPDYSGRLTLMNMLAPAVSPSPGERILFPRSVPSAGGYG 455
Db 373 LSWISPPSPHPSGSQVDLWKIPNYGSSITEATHLAPSVYPPGFGFVLFVFFMKMGPGAYN 432
Qy 456 SGYIDCLIPQWVQHFYQEAAPSQSAVALRVYNPDGTGRNIFEAKLHREGFLTVANCG-- 513
Db 433 ---LPCLLPQBYIISHLASEQAFTVGEAALLHYVDPDGTGRNLGEFKAYPDGFLTCVPNGAS 489
Qy 514 NPVIVPPNGVFRFEANGNOFTYTLAPMGSQGR 547
Db 490 SGPOQLPINGVFRFVSWRSRFPQLKPVGTASSAR 523

RESULT 11
ADC72176
ID ADC72176 standard; protein; 530 AA.
XX
AC ADC72176;
XX
XX 18-DEC-2003 (first entry)
XX
XX Norwalk virus protein 2 amino acid sequence.
XX
XX immune response; non-Norwalk virus agent; immunogen; Norwalk virus;
XX viral pathogen; acute gastroenteritis; virucidal; antiviral vaccine.
XX
XX Norwalk virus.
XX
XX US6572862-B1.
XX
XX
```

PD 03-JUN-2003.
 XX PF 07-JUN-1995; 95US-00486049.
 XX PR 08-NOV-1989; 89US-00433492.
 PR 27-APR-1990; 90US-00515993.
 PR 27-AUG-1990; 90US-00573509.
 PR 06-MAY-1991; 91US-00696454.
 XX PA (BAYU) BAYLOR COLLEGE MEDICINE.
 XX PI Estes MK, Jiang X, Graham DY;
 XX DE WPI; 2003-776005/73.
 DR N-PSDB; ADC72174.
 XX PT Inducing an immune response against non-Norwalk virus agents, comprises
 PT administering an immunogen recombinantly expressed from a cDNA from
 PT Norwalk virus.
 XX PS Example 4; SEQ ID NO 3; 45pp; English.
 XX This invention relates to a novel method of inducing an immune response
 CC in an individual against Norwalk virus and non-Norwalk virus agents, by
 CC orally or parenterally administering an immunogen recombinantly expressed
 CC or synthesised from a cDNA of Norwalk virus given in the specification.
 CC Norwalk virus is one of the most important viral pathogens, causing acute
 CC gastroenteritis. The invention may be used for the development of
 CC compounds with virucidal activity or an antiviral vaccine. The present
 CC sequence is the amino acid sequence of a protein encoded by the Norwalk
 CC virus genome of the invention.
 XX SQ Sequence 530 AA;
 Query Match 39.3%; Score 1141; DB 7; Length 530;
 Best Local Similarity 43.4%; Pred. No. 1.3e-96;
 Matches 249; Conservative 78; Mismatches 169; Indels 78; Gaps 12;
 QY 1 MKMASNDAAPSNDGAA---NLVPEAN-DEVMALEPVVGASIAAPVVGQNIIDPWIRENF 56
 DB 1 MMWASKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSSTAVATAGQVNPDPWIINNF 60
 QY 57 VQAPQGEFTVSPRNSPGEMLLNLLELGPENLYSLRMVNGYAGGMQVQVVLGNAFTA 116
 DB 61 VQAPQGEFTTSPNNTPGDVLDSLGLPHLNPFLHLSQMYNGWGNVRVIMLAGNAFTA 120
 QY 117 GKIIFAAVPPHPFVENISAAQITMCPHVIDVVRQLEPVLPLDPIRNPFFHYNQENTPRM 176
 DB 121 GKIIIVSCIPPGFGSHNLTIAQATLFPHVADVRTLDPIEVLEDRVNLVFNHNDNRNQTM 180
 QY 177 RLVMALYTPLRANSGE-DVFTVSCRVLTRPAPDEFTFLVPPTVESKTKPTLPLTLGE 235
 DB 181 RLVCMLYTPLRGTGGTGSFVAGRVMTCPSPDFNLFLEVPVETVEQKTRPTPLNPLSS 240
 QY 236 LSNRFPAAIDMLYTDNVESTVQVQNGRCTLDTGLTQTTQLVPTQICAFRGTLISQTAR 295
 DB 241 LSNRAPLTPISMGISFDNVQSVQFQNGRCTLDRGLVGTTPVSLSHVAKIRGT----- 293
 QY 296 AADSTDSQRRARNHPLHVQVKNLDGTQYDPTDDIPAVLGAIDFKG-----TVFGVASQ 348
 DB 294 -----SNGTVINLTDLGTTPHPPEG-PAPIGPFDLGGCDWIMHNTQFGHSSQ 340
 QY 349 RDVSGQOQGHYATRAEHAHTDTPKYPKGLTILIKS-GSDDFNTNQPIRFTPVGMGD 407
 DB 341 T-----QYDVTTPDTPFVPHLGSIQANGISGNY-----VG 372
 QY 408 NNWRQ-----WELPDYSGRLTLNNLAPVSPSPGGRILFFRSIVPSAGCYG 455
 DB 373 LSWISPPSPSGQVDLWKIPNYGSSITEATHLAPSVYPPGFGVLELVFFMSKMPGPGAYN 432
 QY 456 SGYIDCLIPQEWQHFQYQEAAPSQSAVALRVYVNPDTGRNITFEAKLHREGFLTIVANGC-- 513
 DB 433 ---LPCLLPQEIYSLHASEQAPTVGEAALLHVDPDGTGRNLGCEFAKYPDGLTLCVPNGAS 489

QY 514 NNPIVVPNGYFRFEANGNQFYTLAPMGSGQGR 547
 DB 490 SGPQQLPINGVFVFWVSWVSRYQLKPVGTASSAR 523
 RESULT 12
 ID AAB49701
 XX AAB49701 standard; protein; 530 AA.
 XX AC AAB49701;
 XX DT 04-APR-2001 (first entry)
 XX Small round structured virus protein SEQ ID 2.
 DE Small round structured virus; SRSV; food poisoning.
 XX Small round structured virus.
 OS Small round structured virus.
 XX WO200079280-A1.
 XX 28-DEC-2000.
 XX 22-JUN-2000; 2000WO-JP004095.
 XX 22-JUN-1999; 99JP-00175928.
 XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.
 XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 WPI; 2001-080848/09.
 DR N-PSDB; AAF29142.
 XX Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX Claim 1; Page 42-45; 84pp; Japanese.
 PS This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX SQ Sequence 530 AA;
 Query Match 38.9%; Score 1131; DB 4; Length 530;
 Best Local Similarity 43.2%; Pred. No. 1.1e-95;
 Matches 248; Conservative 77; Mismatches 171; Indels 78; Gaps 12;
 QY 1 MKMASNDAAPSNDGAA---NLVPEAN-DEVMALEPVVGASIAAPVVGQNIIDPWIRENF 56
 DB 1 MMWASKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSSTAVATAGQVNPDPWIINNF 60
 QY 57 VQAPQGEFTVSPRNSPGEMLLNLLELGPENLYSLRMVNGYAGGMQVQVVLGNAFTA 116
 DB 61 VQAPQGEFTTSPNNTPGDVLDSLGLPHLNPFLHLSQMYNGWGNVRVIMLAGNAFTA 120
 QY 117 GKIIFAAVPPHPFVENISAAQITMCPHVIDVVRQLEPVLPLDPIRNPFFHYNQENTPRM 176
 DB 121 GKIIIVSCIPPGFGSHNLTIAQATLFPHVADVRTLDPIEVLEDRVNLVFNHNDNRNQTM 180
 QY 177 RLVMALYTPLRANSGE-DVFTVSCRVLTRPAPDEFTFLVPPTVESKTKPTLPLTLGE 235
 DB 181 RLVCMLYTPLRGTGGTGSFVAGRVMTCPSPDFNLFLEVPVETVEQKTRPTPLNPLSS 240
 QY 236 LSNRFPAAIDMLYTDNVESTVQVQNGRCTLDTGLTQTTQLVPTQICAFRGTLISQTAR 295

[illegible]

RESULT 13
AAB49703
ID AAB49703 standard; protein; 544 AA.
XX
AAB49703;
XX AC
XX DT 04-APR-2001 (first entry)
XX
XX Small round structured virus protein SEQ ID 4.
XX DE
XX Small round structured virus; SRSV; food poisoning.
XX KW
XX Small round structured virus.
XX OS
XX Small round structured virus.
XX WO200079280-A1.
XX PN
XX 28-DEC-2000.
XX PD
XX 22-JUN-2000; 2000WO-JP004095.
XX PF
XX 22-JUN-1999; 95JP-00175928.
XX PR
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX PA (DENK-) DENKA SEIKEN KK.
XX PA
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato S;
XX WIPI; 2001-080848/09.
XX DR N-PSDB: AAF29144.
XX DR

XX	Kit for the detection and typing of small round-structured virus (SRSV)
PT	strains for investigation of food poisoning outbreaks, contains
PT	antibodies.
XX	
XX	Claim 1; Page 47-49; 84pp; Japanese.
XX	
CC	This invention relates to a kit for the detection and typing of small
CC	round structured virus (SRSV) strains. The kit contains antibodies
CC	directed against peptides represented in sequences AAB49700 - AAB49710,
CC	which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC	AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC	used for detecting and typing strains of SRSV in order to prevent the
CC	spread of infection and to examine the epidemiology of outbreaks
XX	
XX	Sequence 544 AA;
SQ	
Query Match	38.7%; Score 1124.5; DB 4; Length 544;
Best Local Similarity	44.2%; Pred. No. 4.5e-95;
Matches 250; Conservative	80; Mismatches 132; Indels 43; Gaps 14

[illegible]

RESULT 14	
AA57091	
ID	AA57091 standard; protein; 530 AA.
XX	
AC	AA57091;
XX	
DT	27-AUG-2003 (revised)
DT	25-MAR-2003 (revised)
DT	05-OCT-1994 (first entry)
XX	
DE	Small round virus SRSV/KY/89 capsid protein.
XX	
KW	pathogen; acute gastroenteritis; food poison
KW	diagnostic assay; human calicivirus; small ro
KW	Norwalk virus; capsid protein.
XX	
OS	Small round structured virus.
XX	
PN	WO9405700-A2.
XX	
PD	17-MAR-1994.
XX	
XX	07-SEP-1993; 93WO-US008447.
XX	
PR	07-SEP-1992; 92US-00941365.
XX	
PA	(BAYU) BAYLOR COLLEGE MEDICINE.
XX	
PI	Matson DO, Estes MK, Jiang X, Graham DY;
XX	
DR	WPI; 1994-101125/12.

Wed Jun 2 09:13:24 2004

DR N-PSDB; AAQ56832.
 XX DNA from Norwalk and related viruses - used for preparing prods. for use
 PT in diagnostic assays, detection and vaccines for Norwalk and related
 PT viruses.
 XX Example 7; Fig 13a; 156pp; English.
 XX
 XX The known sequence for Norwalk virus was used to obtain the sequence of
 CC other Norwalk-related viruses such as SRSV/KY/89, an agent from a stool
 CC from an outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide
 CC cDNA sequence includes part of the polymerase region and the capsid
 CC region of the genome; the deduced amino acid sequences are AAR57052 and
 CC AAR57091, respectively. Expression of fragments and derivs. of Norwalk-
 CC related viruses permits development of diagnostic assays to detect
 CC antibodies, antigens, viral genetic material or antivirals. (Updated on
 CC 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS
 CC field.)
 XX
 XX Sequence 530 AA;
 SQ
 Query Match 37.7%; Score 1097; DB 2; Length 530;
 Best Local Similarity 42.8%; Pred. No. 1.6e-92;
 Matches 246; Conservative 78; Mismatches 171; Indels 80; Gaps 14;
 QY 1 MKMASNDAAPSNDGAA--NLVPEAN-DEVMALEPVVGASIAAPVVGQNIIDPWIRENF 56
 DB 1 MMWASKDATTSSVDGASASQVLPEVNASDPLAMDVPAGSSTAVATAGQVNPIDWIINF 60
 QY 57 VQAPQGEFTVSPRNSPGEMLNLLELGPENLPYLSHLSRMVNGYAGGMQVQVVLGNAFTA 116
 DB 61 VQAPQGEFTISPNNTPGDVLFDSLGLPHLPFLSHLSQMYNGVGNMRVRLAGNAFSA 120
 QY 117 GKLIFAAVPFPFVENISAAQITMCPHVIVDVRQLEPVLPLPDIRNRFPHYNOENTPRM 176
 DB 121 GKLIIVCCIPFGSGQQITIAQATLFPHVIAADVRLDPIEVPLEDVRNVLFNHNRNQTM 180
 QY 177 RLVALMYTPLRANGS--DVTFTVSCRVLTRPADPFEFTFLVPPTVESKTKPFTLPLTL 235
 DB 181 RLVCMLYTPLSTGGTGDSPFVAGRVNMTCPSPDNFLFLVPFTVEQKTRPFTLPNPLSS 240
 QY 236 LNSRFPAAIDMLYTDPNESIVVQPNQGRCTLDGTTLQGTTLVPTQICAFRGTLSQSTAR 295
 DB 241 LNSRFAPLPSGMISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGT----- 293
 QY 296 AADSTDSQPRARNHPLHVQVKNLDGTQYDPTDDIPAVLGALDFKG-----TVFGVASQ 348
 DB 294 -----SNGTVINLTDLGTPFHPFEG-PAPIGFPDLGGCDWHINMTQFGHSSQ 340
 QY 349 RDVSGQQQGHYATRAHEAHIIDTDPKYAPKLGITLIKSGSDFFNTNQPIRFTPVGMGD 407
 DB 341 T-----QYDVTDTTDSVPHLGSIQANGIGSGNY-----IGV 372
 QY 408 NNWRQ-----WELPDYSGRLTFLNNMLAPAV-SPSPGGERILFRSIVPSAGGY 454
 DB 373 LSWVSPSPHSPSGQVDLWKIPNYGSSITEATHLAPSVVSPGF-GEVLVFFMSKIPGPG- 430
 QY 455 GSGYIDCLIPQBWQHFQEAAPSQSAVALRVYVNPDTGRNIEAKLHREGFLIVACNG- 513
 DB 431 --DSLPCLLPOGIYSHLASEQAPTVGEGLLHYVDPDTRNLGZEFKAYPDGFLTVCVNGA 488
 QY 514 -NNPIVVPNGYFRFEACWGNQFYTLAPMGSGQGR 547
 DB 489 SSGPQOLPINGVVFVSVWSRYQLKPVGTASTAR 523
 RESULT 15
 AAB49702
 ID AAB49702 standard; protein; 546 AA.
 XX
 AC AAB49702;
 XX
 DT 04-APR-2001 (first entry)

XX Small round structured virus protein SEQ ID 3.
 XX Small round structured virus; SRSV; food poisoning.
 XX Small round structured virus.
 XX WO200079280-A1.
 XX
 XX 28-DEC-2000.
 XX
 XX 22-JUN-2000; 2000WO-JP004095.
 XX
 XX 22-JUN-1999; 99JP-00175928.
 XX
 XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 XX (DENK-) DENKA SEIKEN KK.
 XX
 XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 XX
 XX WPI: 2001-080848/09.
 XX N-PSDB; AAP29143.
 XX
 XX Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX
 XX Claim 1; Page 45-47; 84pp; Japanese.
 XX
 XX This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX
 XX Sequence 546 AA;
 SQ
 Query Match 37.5%; Score 1090.5; DB 4; Length 546;
 Best Local Similarity 42.6%; Pred. No. 6.6e-92;
 Matches 250; Conservative 75; Mismatches 173; Indels 89; Gaps 16;
 QY 1 MKMASNDAAPSNDGAA--NLVPEAND-EVMALEPVVGASIAAPVVGQNIIDPWIRENF 56
 DB 1 MMWASKDAPQASDAGSAGQVPEVNTADPLPEPVAGPTTAVATAGQVNMIDPWVNF 60
 QY 57 VQAPQGEFTVSPRNSPGEMLNLLELGPENLPYLSHLSRMVNGYAGGMQVQVVLGNAFTA 116
 DB 61 VQAPQGEFTISPNNTPGDVLFDSLGLPHLPFLSHLSQMYNGVGNMRVRLAGNAFSA 120
 QY 117 GKLIFAAVPFPFVENISAAQITMCPHVIVDVRQLEPVLPLPDIRNRFPHYNOENTPRM 176
 DB 121 GKLIIVCCIPFGTSSUTIAQATLFPHVIAADVRLDPIEVPLEDVRNVLVHTN-DNQPTM 179
 QY 177 RLVALMYTPLRANGS--EDVFTVSCRVLTRPADPFEFTFLVPPTVESKTKPFTLPLTL 233
 DB 180 RLVCMLYTPLRTGGSGNSDSFVAVAGVLTAPSDFSFLVLPPTIEQKTRAFVPMPL 239
 QY 234 GELNSRFPAAIDMLYTDPNESIVVQPNQGRCTLDGTTLQGTTLVPTQICAFRGTLSQST 293
 DB 240 QTLNSRFPFSLTQGMILSPDASQVQVQNGRCLLDGQLLGTTPATSGQLFVRCK-INQG 298
 QY 294 ARAADSTDSQPRARNHPLHVQVKNLDGTQYDPTDDIPAVLGALDF-----KG 340
 DB 299 ARTLNLT-----VDGKPFMAFDS-PAPVGFDPFGKCDWHMRISKTPNN 341
 QY 341 TVFG-----VASQDVSGQEQGHYATRAHEAHIIDTDPKYAPKLGITLIKSGSDFFN-- 393
 DB 342 TSSGDPMRGSVQTNVQG-----FVPHLGSIQF---DEVFNHP 376
 QY 394 -----TNQPIRFTPVGMGNWRQWELPDYSGRILTLMNMLAPAVSPSPGGERILF 443

Db 377 TGDYIGTIEWISQP--STPPGTDN---LWEIPDYGSSLQAANLAPPVPPPGGEALVY 431
Qy 444 FRSIVPSAGGYS-GYIDCLIPQEWVQHFOEAAFPQSQAVALVRYVNPDTGRNIFEAKLH 502
Db 432 FVSAPFGPNRSAPNDVPCLLPQEVITHFVSEQAFTMGDAALLHYVDPDTNRNLGEFKLY 491
Qy 503 REGFLTVA--NCGNNPIVVPPNGYFRFEAWGNQFYTLAPMGSGQRR 547
Db 492 PGGYLTCTVPNGVGAGPQQLPLNGVFLFVSWVSRYQLKXPGVTASTAR 538

Search completed: June 1, 2004, 13:46:01
Job time : 49.2483 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 13.104 Seconds
(without alignments)
2166.837 Million cell updates/sec

Title: US-09-926-799-10
Perfect score: 2306
Sequence: 1 MKWASNDAAFSNDGAAALNP.....GNQPYTLAPMGSGGRRRAQ 550

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1141	39.3	530	4	US-08-486-049-3
2	280.5	9.7	668	4	US-08-617-594A-4
3	277.5	9.5	626	4	US-09-590-020-7
4	276.5	9.5	622	4	US-09-590-020-6
5	276.5	9.5	623	4	US-09-590-020-2
6	276.5	9.5	623	4	US-09-590-020-4
7	272.5	9.4	669	4	US-08-617-594A-2
8	129.5	4.5	2318	3	US-08-091-219-24
9	129.5	4.5	2318	4	US-09-660-541-24
10	119	4.1	2647	2	US-08-583-562B-8
11	119	4.1	2647	2	US-08-779-113-8
12	109.5	3.8	1079	2	US-08-929-967-8
13	109.5	3.8	1737	4	US-09-309-572-13
14	109.5	3.8	1737	4	US-08-718-096-13
15	102.5	3.5	335	4	US-09-252-991A-28530
16	102	3.5	571	4	US-09-134-001C-3865
17	100	3.4	2232	3	US-09-091-219-25
18	100	3.4	2232	4	US-09-660-541-25
19	100	3.4	2247	3	US-09-091-219-2
20	100	3.4	2247	4	US-09-660-541-2
21	99	3.4	1091	6	US-08-475-886-2
22	99	3.4	2227	3	US-08-475-886-4
23	99	3.4	2227	3	US-08-475-886-6
24	99	3.4	2227	3	US-08-475-886-5
25	99	3.4	2227	3	US-08-397-232-2
26	99	3.4	2227	3	US-08-397-232-4
27	99	3.4	2227	3	US-09-171-387-2

28	99	3.4	2227	4	US-09-653-499-2	Sequence 2, Appli
29	99	3.4	2227	4	US-09-653-499-4	Sequence 4, Appli
30	99	3.4	2227	4	US-09-653-499-6	Sequence 6, Appli
31	99	3.4	2227	4	US-10-104-966-12	Sequence 12, Appl
32	99	3.4	2227	4	US-10-135-988-2	Sequence 2, Appli
33	99	3.4	2227	4	US-10-135-988-4	Sequence 4, Appli
34	99	3.4	2227	4	US-10-135-988-6	Sequence 6, Appli
35	98.5	3.4	1257	2	US-08-750-152A-2	Sequence 2, Appli
36	98	3.4	226	3	US-09-091-219-5	Sequence 5, Appli
37	98	3.4	226	4	US-09-660-541-5	Sequence 5, Appli
38	97.5	3.4	713	2	US-08-849-212-4	Sequence 4, Appli
39	97.5	3.4	725	1	US-08-448-170-4	Sequence 4, Appli
40	97.5	3.4	725	3	US-08-961-803-7	Sequence 7, Appli
41	97.5	3.4	832	4	US-09-489-039A-12438	Sequence 12438, A
42	97	3.3	839	1	US-08-087-016-2	Sequence 2, Appli
43	97	3.3	1269	4	US-09-645-456A-15	Sequence 15, Appl
44	97	3.3	1269	4	US-09-425-324A-15	Sequence 15, Appl
45	97	3.3	1269	4	US-09-645-791-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-08-486-049-3
; Sequence 3, Application US/08486049
; Patent No. 6572862
; GENERAL INFORMATION:
; APPLICANT: Estes, Mary K
; APPLICANT: Jiang, Xi
; APPLICANT: Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Ave., N.W.
; CITY: Washington, D.C.
; STATE:
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/486,049
; FILING DATE: June 7, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311.023
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-486-049-3

Query Match 39.3%; Score 1141; DB 4; Length 530;
Best Local Similarity 43.4%; Pred. No. 2.4e-106;
Matches 249; Conservative 78; Mismatches 169; Indels 78; Gaps 12;
Qy 1 MKWASNDAAFSNDGAA---NLVPEAN-DEVWALEPVVVGASIAAPVVGQNIIDPWIRENF 56
Db 1 MMWASKDATSSVDGASGAGQLVPEVNASDPLAMDPLVDGSSSTAVATAGOVNPDWIIINF 60

QY 57 VOAPQGEFTVSPRNSPGEMLLNLELGPENLPYLHSLRMVYAGGQVQVVLGNAGFTA 116
DB 61 VOAPQGEFTVSPRNSPGEMLLNLELGPENLPYLHSLRMVYAGGQVQVVLGNAGFTA 120
QY 117 GKIIIFAAVPPHPPVENISAAQITMCPHVIVDVROLEPVLPLDIPRNRFFHYNQENTPRM 176
DB 121 GKIIIVSCIPPFGSHNLTAQATLPHVIADVTLDPIEVLDEVRNLPHNDRNQOTM 180
QY 177 RLVMAMLYTPLRANGSE-DVFTVSCRVLTRPAPDEFTFLVPPTVESKTKPFTLPILTGE 235
DB 181 RLVMAMLYTPLRANGSE-DVFTVSCRVLTRPAPDEFTFLVPPTVESKTKPFTLPILTGE 240
QY 236 LNSNRPRAIDMLYTDPNESIVVQPNQGRCTLDGTQVDTDDIPAVLGADFKG-----TVFGVASQ 348
DB 241 LNSNRPRAIDMLYTDPNESIVVQPNQGRCTLDGTQVDTDDIPAVLGADFKG-----TVFGVASQ 348
QY 296 AADSTDSQARARNHPLHVQVKNLDTQVDTDDIPAVLGADFKG-----TVFGVASQ 348
DB 294 AADSTDSQARARNHPLHVQVKNLDTQVDTDDIPAVLGADFKG-----TVFGVASQ 348
QY 349 RNVSGQEQGHVATRAHEAHIIDTTPDKYAPKLGTILIKS-GSDDPNTNQIRFTPVGMGD 407
DB 341 T-----QVDTTITDTPVPHLGSIOANGISGNY-----VGV 372
QY 408 NNWRO-----WELPDYSGLRLTNMMLAPAVSPSPGGERILFRSIVPSAGYG 455
DB 373 LSWISPPSHPSGSQVDLWIKIPNGSSITATHLAPSVYPPGGEVLVFMNKMPPGAGYN 432
QY 456 SGVIDCLIQEWOHYQEAAPSQSAVALRVYVNDTGNRIPEAKLHREGFLTVANCG-- 513
DB 433 ---LPCLLQEYISHLASQAPTVEAALHVDVDTGRLNLFGEKAYPDGFLTCVPNGAS 489
QY 514 NNPIVPPNGYFRAEAGNQFTYTLAPMGSGQGR 547
DB 490 SGFQQLPINGVFFVSVMSRFFYQLKPVGTASSAR 523
RESULT 2
US-09-617-594A-4
; Sequence 4, Application US/09617594A
; Patent No. 6541458
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.1
; CURRENT APPLICATION NUMBER: US/09/617,594A
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: 2000-03-30
; PRIOR FILING DATE: 2000-02-11
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-617-594A-4
Query Match 9.7%; Score 280.5; DB 4; Length 668;
Best Local Similarity 25.7%; Pred. No. 3.5e-19;
Matches 103; Conservative 56; Mismatches 147; Indels 95; Gaps 17;
QY 11 SNDGAANLVPEANDEV--MALEPVVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSP 68
DB 125 ADDGSVT-TPEQTLVGGVIAEPNAQMSAVADVATGKSDVSEW--EAFPSF-HTSVNWS 180
QY 69 RNSPCEMLNLELGPENLPYLHSLRMVYAGGQVQVVLGNAGFTAAGKIIFAAVPPHF 128
DB 181 SETQKILFKQSLGFLNLPYLHSLRMVYAGGQVQVVLGNAGFTAAGKIIFAAVPPHF 240

QY 129 -PVENISAAQITMCPHVIVDVROLEPVLPLDIPRNRFFHYNQENTPRMELVAMLYTPL- 186
DB 241 DEVQSTMLQY---PHVLFDAQVPEVIFTPDLRNSLYHL-MSDTRTSLVIMVYNDLI 296
QY 187 -----RANSGBEDVFTVSCRVLTRPAPDEFTFLVPPTVESKTKPFTLPILTGLSLSR 241
DB 297 NPYANDSNSSGCIIVTVE---TKGPDQFKHLKPPG-----SMLTHGSIPLD 342
QY 242 PAADIMLYTDPNES-----IVVQPNQGRCTLDGTQVDTDDIPAVLGADFKG-----FVQANRHF 374
DB 343 PKSSSLWTGNRHSWIDTDFVIK-----FVQANRHF 374
QY 298 D---STDSQARARNHPLHVQVKNLDTQVDTDDIPAVLGADFKG-----FVQANRHF 374
DB 375 DFNQETAGWSTFRPRPIITVSEKGSK-----LG-----IGVATDSIVFG- 415
QY 355 QEGHYATRAHEAHIIDTTPDKYAPKLGTILIKS-GSDDPNTNQIRFTPVGMGD 407
DB 416 -----IPDGWPDPTTPEKLTIPAGDYAITNGNNDITT 447
RESULT 3
US-09-590-020-7
; Sequence 7, Application US/095900020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnius, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590, 020
; CURRENT FILING DATE: 2000-06-08
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 7
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-7
Query Match 9.5%; Score 277.5; DB 4; Length 626;
Best Local Similarity 23.7%; Pred. No. 6.2e-19;
Matches 127; Conservative 79; Mismatches 209; Indels 121; Gaps 23;
QY 11 SNDGAANLVPEANDEV--MALEPVVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSP 68
DB 80 ADDGSIT-APEQGTWVGVIATPESQMSAAMATGKSDVSEW--EAFPSF-HTSVNWS 135
QY 69 RNSPCEMLNLELGPENLPYLHSLRMVYAGGQVQVVLGNAGFTAAGKIIFAAVPPHF 128
DB 136 SETQKILFKQSLGFLNLPYLHSLRMVYAGGQVQVVLGNAGFTAAGKIIFAAVPPHF 195
QY 129 -PVENISAAQITMCPHVIVDVROLEPVLPLDIPRNRFFHYNQENTPRMELVAMLYTPLR 187
DB 196 DFNQSTMLQY---PHVLFDAQVPEVIFTPDLRNSLYHL-MSDTRTSLVIMVYNDLI 251
QY 188 ANSGEDVFTVSC--RVLTRPAPDEFTFLVPPTVESKTKPFTLPILTGLSLSRFPAAI 245
DB 252 NPYANDTNSSGCIIVTVEKPGDPDFKHLKPPG-----SMLTHGSIPLDIPKSS 301
QY 246 DMLYTDPNES-----IVVQPNQGRCTLDGTQVDTDDIPAVLGADFKG-----FVQANRHF 374
DB 302 SLWIGNRHSWIDTDFIIRP-----FVQANRHF 374
QY 299 STDSQARARNHPLHVQVKNLDTQVDTDDIPAVLGADFKG-----FVQANRHF 374
DB 334 ETAGWSTFRPRPIITVSEKGSK-----AKLG-----IGVATDSIVFG- 415


```
Qy 359 HYATRAHEAHIDTDPKYPAPKLILKSGS--DDFNTNQPIRFTPVGMGDNWNRQWELPD 417
Db 371 -----IPDGWPDPTTGPGLIPAGDYAITNGTGNDDITATGYDTADIIKNNTNR----- 419
Qy 418 YSGRLTLNMNLAPAVSPSPFCERILFFRSIVPSA---GGYSGYIDC-LIPOEWV----- 468
Db 420 -----GMYICGSLQRAW-GDKKISNTAFITTTATLDGDNKKINPCNTIDQSKIVVFQD 471
Qy 469 OHFYQEAAPSQSAVALVRY-----VNPDTGRNIFEAKLHREGFLTVANGNNPI 517
Db 472 AHVGKKAQTSDDTLALLGYTGIGRQAIGSDRDRVVRISTLPETG-----ARGGNHPI 523

RESULT 4
US-09-590-020-6
; Sequence 6, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilinis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-6

Query Match          9.5%; Score 276.5; DB 4; Length 622;
Best Local Similarity 31.1%; Pred. No. 7.8e-19;
Matches 73; Conservative 39; Mismatches 102; Indels 21; Gaps 8;

Qy 12 NDGAANLVPEANDEV--MALEPVVVGASIAAPVVGQQNIIDPWIRENFVQAPQGFTVSPR 69
Db 81 DDGSIT-TPEQGTWVGVIAPSPSAQMSAAADMATGKSVDSSEW--EAPFSF-HTSVNNWSTS 136
Qy 70 NSPGEMLLNLELGPENLPYLHSLRMYNGVAGGMQVQVVLGNAGTAFTAGKIIFAAVPPHPP 129
Db 137 ETQKILFKQSLGPLLNPYLHSLAKLYVAMSGSVEVRFSGSGVFGKLAIVVPP--G 194
Qy 130 VENISAAQITMCPHVIVDVRLPVLPLDIRNRFHYNQENTPRMLVAMLYTPLRAN 189
Db 195 IEPVQSTSMQLQPHVLFDAQVEPVIFAIPDLRSNLVHL-MSDTDTTSLVIMVYNDLINP 253
Qy 190 SGEDVFTVSC--RVLTRPADPDEFTFLVPPTVESKTKPFTLPILTIGELNSRPP 242
Db 254 YANDTNSSGCIIVTETKPGDPDFKHLKPPG-----SMLTHGSGVPSDLIP 298

RESULT 5
US-09-590-020-2
; Sequence 2, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilinis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9

Qy 12 NDGAANLVPEANDEV--MALEPVVVGASIAAPVVGQQNIIDPWIRENFVQAPQGFTVSPR 69
Db 81 DDGSIT-TPEQGTWVGVIAPSPSAQMSAAADMATGKSVDSSEW--EAPFSF-HTSVNNWSTS 136
Qy 70 NSPGEMLLNLELGPENLPYLHSLRMYNGVAGGMQVQVVLGNAGTAFTAGKIIFAAVPPHPP 129
Db 137 ETQKILFKQSLGPLLNPYLHSLAKLYVAMSGSVEVRFSGSGVFGKLAIVVPP--G 194
Qy 130 VENISAAQITMCPHVIVDVRLPVLPLDIRNRFHYNQENTPRMLVAMLYTPLRAN 189
Db 195 IEPVQSTSMQLQPHVLFDAQVEPVIFAIPDLRSNLVHL-MSDTDTTSLVIMVYNDLINP 253
Qy 190 SGEDVFTVSC--RVLTRPADPDEFTFLVPPTVESKTKPFTLPILTIGELNSRPP 242
Db 254 YANDTNSSGCIIVTETKPGDPDFKHLKPPG-----SMLTHGSGVPSDLIP 298

RESULT 6
US-09-590-020-4
; Sequence 4, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilinis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-4

Query Match          9.5%; Score 276.5; DB 4; Length 623;
Best Local Similarity 30.6%; Pred. No. 7.8e-19;
Matches 72; Conservative 41; Mismatches 101; Indels 21; Gaps 8;

Qy 12 NDGAANLVPEANDEV--MALEPVVVGASIAAPVVGQQNIIDPWIRENFVQAPQGFTVSPR 69
Db 81 DDGSIT-TPEQGTWVGVIAPSPSAQMSAAADMATGKSVDSSEW--EAPFSF-HTSVNNWSTS 136
Qy 70 NSPGEMLLNLELGPENLPYLHSLRMYNGVAGGMQVQVVLGNAGTAFTAGKIIFAAVPPHPP 129
Db 137 ETQKILFKQSLGPLLNPYLHSLAKLYVAMSGSVEVRFSGSGVFGKLAIVVPP--G 194
Qy 130 VENISAAQITMCPHVIVDVRLPVLPLDIRNRFHYNQENTPRMLVAMLYTPLRAN 189
Db 195 IEPVQSTSMQLQPHVLFDAQVEPVIFAIPDLRSNLVHL-MSDTDTTSLVIMVYNDLINP 253
Qy 190 SGEDVFTVSC--RVLTRPADPDEFTFLVPPTVESKTKPFTLPILTIGELNSRPP 242
Db 254 YANDTNSSGCIIVTETKPGDPDFKHLKPPG-----SMLTHGSGVPSDLIP 298

RESULT 7
US-09-590-020-5
; Sequence 5, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilinis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-2

Query Match          9.5%; Score 276.5; DB 4; Length 623;
Best Local Similarity 30.6%; Pred. No. 7.8e-19;
Matches 72; Conservative 41; Mismatches 101; Indels 21; Gaps 8;

Qy 12 NDGAANLVPEANDEV--MALEPVVVGASIAAPVVGQQNIIDPWIRENFVQAPQGFTVSPR 69
Db 81 DDGSIT-TPEQGTWVGVIAPSPSAQMSAAADMATGKSVDSSEW--EAPFSF-HTSVNNWSTS 136
Qy 70 NSPGEMLLNLELGPENLPYLHSLRMYNGVAGGMQVQVVLGNAGTAFTAGKIIFAAVPPHPP 129
Db 137 ETQKILFKQSLGPLLNPYLHSLAKLYVAMSGSVEVRFSGSGVFGKLAIVVPP--G 194
Qy 130 VENISAAQITMCPHVIVDVRLPVLPLDIRNRFHYNQENTPRMLVAMLYTPLRAN 189
Db 195 IEPVQSTSMQLQPHVLFDAQVEPVIFAIPDLRSNLVHL-MSDTDTTSLVIMVYNDLINP 253
Qy 190 SGEDVFTVSC--RVLTRPADPDEFTFLVPPTVESKTKPFTLPILTIGELNSRPP 242
Db 254 YANDTNSSGCIIVTETKPGDPDFKHLKPPG-----SMLTHGSGVPSDLIP 298

RESULT 6
US-09-590-020-4
; Sequence 4, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilinis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-4

Query Match          9.5%; Score 276.5; DB 4; Length 623;
Best Local Similarity 30.6%; Pred. No. 7.8e-19;
Matches 72; Conservative 41; Mismatches 101; Indels 21; Gaps 8;

Qy 12 NDGAANLVPEANDEV--MALEPVVVGASIAAPVVGQQNIIDPWIRENFVQAPQGFTVSPR 69
Db 81 DDGSIT-TPEQGTWVGVIAPSPSAQMSAAADMATGKSVDSSEW--EAPFSF-HTSVNNWSTS 136
Qy 70 NSPGEMLLNLELGPENLPYLHSLRMYNGVAGGMQVQVVLGNAGTAFTAGKIIFAAVPPHPP 129
Db 137 ETQKILFKQSLGPLLNPYLHSLAKLYVAMSGSVEVRFSGSGVFGKLAIVVPP--G 194
Qy 130 VENISAAQITMCPHVIVDVRLPVLPLDIRNRFHYNQENTPRMLVAMLYTPLRAN 189
Db 195 IEPVQSTSMQLQPHVLFDAQVEPVIFAIPDLRSNLVHL-MSDTDTTSLVIMVYNDLINP 253
Qy 190 SGEDVFTVSC--RVLTRPADPDEFTFLVPPTVESKTKPFTLPILTIGELNSRPP 242
Db 254 YANDTNSSGCIIVTETKPGDPDFKHLKPPG-----SMLTHGSGVPSDLIP 298

RESULT 7
US-09-590-020-5
; Sequence 5, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilinis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
```

US-09-617-594A-2
 ; Sequence 2, Application US/09617594A
 ; Patent No. 6541458
 ; GENERAL INFORMATION:
 ; APPLICANT: Audonnet, et al.
 ; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V
 ; FILE REFERENCE: 454313-3151.1
 ; CURRENT APPLICATION NUMBER: US/09/617,594A
 ; CURRENT FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: 60/193,332
 ; PRIOR FILING DATE: 2000-03-30
 ; PRIOR APPLICATION NUMBER: France 00 01761
 ; PRIOR FILING DATE: 2000-02-11
 ; PRIOR APPLICATION NUMBER: France 99 09421
 ; PRIOR FILING DATE: 1999-07-16
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 669
 ; TYPE: PRT
 ; ORGANISM: Feline calicivirus
 ; US-09-617-594A-2

Query Match 9.4%; Score 272.5; DB 4; Length 669;
 Best Local Similarity 30.2%; Pred. No. 2.2e-18;
 Matches 79; Conservative 43; Mismatches 105; Indels 35; Gaps 11;
 QY 13 DGANL-VPEANDEV--MALEPVVGASIAAPVVGQONIDPWIENFVQAPQGEFTVSPR 69
 DB 126 DGSSITTEQGTFLGGVTAESQAQWATAADATGKSDVSEW--BSFFSF-HTSVNWSTS 182
 QY 70 NSPGEMLLNLELGPENLYSLHSLRMVNGYAGQMVOVVLGNAFTAGKIIFAAVPPHF- 128
 DB 183 ETQKILFKQSLPLNLYSLHSLKLVAVSGSDVRFSGSGVFGKLAIVVPPGVD 242
 QY 129 PVENISAAQITMCPHVIVDRQLEPVLPLPDIRNRFFHYNQENTPRMLVAMLYTPL-- 186
 DB 243 PVQSTSMLOQ---PHVLFARQVPEVIFSPIDURSTLYHL-MSDITTTSLVIMVNDLIN 298
 QY 187 ----RANGSDVFTVSCRVLTRPADPFEFTFLVPPTVESKTKPFTLPILTLGELNSRPP 242
 DB 299 PYANDSNSSCIVTVE---TKGPDFKFHLKPPG-----SMLTHGSIPSDLIP 344
 QY 243 AA----IDMLYTPDNESIVVQP 260
 DB 345 KSSSLWIGNRYWSDITDFVIRP 366

RESULT 8
 US-09-091-219-24
 ; Sequence 24, Application US/09091219
 ; Patent No. 6171592
 ; GENERAL INFORMATION:
 ; APPLICANT: STUDDERT, Michael J.
 ; APPLICANT: CRABB, Brendan S.
 ; APPLICANT: FENG, Li
 ; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
 ; FILE REFERENCE: 040268/0151
 ; CURRENT APPLICATION NUMBER: US/09/091,219
 ; CURRENT FILING DATE: 1998-10-05
 ; EARLIER APPLICATION NUMBER: PCT/AU96/00815
 ; EARLIER FILING DATE: 1996-12-18
 ; EARLIER APPLICATION NUMBER: AU PN7201
 ; EARLIER FILING DATE: 1995-12-18
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 24
 ; LENGTH: 2318
 ; TYPE: PRT
 ; ORGANISM: Foot-and-mouth disease virus
 ; US-09-091-219-24
 Query Match 4.5%; Score 129.5; DB 3; Length 2318;

Best Local Similarity 21.2%; Pred. No. 0.0051;
 Matches 75; Conservative 35; Mismatches 131; Indels 113; Gaps 15;
 QY 79 LELGPELNLYSLHSLRMVNGYAGQMVOVVLGNAFTAGKIIFAAVPPHFVENISAAQI 138
 DB 367 LELPTDHHKGVVGLTDSYAYMRNGWDVEVTAVGNQFNGGCLLVAMVPFLYSIQKRELYQL 426
 QY 139 TMCCHVIVDRQLEPVLPLPDIR-NRFFHYNQENTPRMLVAMLYTPLRANGSDVFTV 197
 DB 427 TLFPHQFINPRNTMTAHITVPFVGNRYDQYKHKP--WTLVVMVAPLTVN----- 476
 QY 198 SCRVLTRPADPFEFTFLVPPTVESKTKPFTLPILTLGELNSR--FPAADID-----MLYT 250
 DB 477 -----TEGAPQIKVYANIAPT-----NHHVAGEFFPSKEGIFPVACSDGYGGLVTT 521
 QY 251 DPNES-----IVVQPN---GRCTLDGTGTTQTLVPTQICAFRGTLISQTAARAADSTD 301
 DB 522 DPKTADPVYKGVNFPNQLPGRFT---NLLDVAEACPT-FLRFEggVpYVTTK---TD 573
 QY 302 SPQARNHPLHVQVKNLDG-----TQYD-----PTD----- 327
 DB 574 SDRLAQFDMSLAAKQMSNTFLAGLAQYTYQVSGTINLHFMFTGPTDAKARYMVAYAPPG 633
 QY 328 -----DIPAVLGAIIDFKGTGFGVASQSDVSG 353
 DB 634 MBPPKTPFAAAHCHIAEMWDTGLNSKFTFSIP-YLSAADYAYTASGVAETTNVQG 686

RESULT 9
 US-09-660-541-24
 ; Sequence 24, Application US/09660541
 ; Patent No. 6531136
 ; GENERAL INFORMATION:
 ; APPLICANT: STUDDERT, Michael J.
 ; APPLICANT: CRABB, Brendan S.
 ; APPLICANT: FENG, Li
 ; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
 ; FILE REFERENCE: 040268/0151
 ; CURRENT APPLICATION NUMBER: US/09/660,541
 ; CURRENT FILING DATE: 2000-09-12
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/091,219
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-05
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: AU PN7201
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1995-12-18
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 24
 ; LENGTH: 2318
 ; TYPE: PRT
 ; ORGANISM: Foot-and-mouth disease virus
 ; US-09-660-541-24
 Query Match 4.5%; Score 129.5; DB 4; Length 2318;
 Best Local Similarity 21.2%; Pred. No. 0.0051;
 Matches 75; Conservative 35; Mismatches 131; Indels 113; Gaps 15;
 QY 79 LELGPELNLYSLHSLRMVNGYAGQMVOVVLGNAFTAGKIIFAAVPPHFVENISAAQI 138
 DB 367 LELPTDHHKGVVGLTDSYAYMRNGWDVEVTAVGNQFNGGCLLVAMVPFLYSIQKRELYQL 426
 QY 139 TMCCHVIVDRQLEPVLPLPDIR-NRFFHYNQENTPRMLVAMLYTPLRANGSDVFTV 197
 DB 427 TLFPHQFINPRNTMTAHITVPFVGNRYDQYKHKP--WTLVVMVAPLTVN----- 476
 QY 198 SCRVLTRPADPFEFTFLVPPTVESKTKPFTLPILTLGELNSR--FPAADID-----MLYT 250
 DB 477 -----TEGAPQIKVYANIAPT-----NHHVAGEFFPSKEGIFPVACSDGYGGLVTT 521
 QY 251 DPNES-----IVVQPN---GRCTLDGTGTTQTLVPTQICAFRGTLISQTAARAADSTD 301
 DB 522 DPKTADPVYKGVNFPNQLPGRFT---NLLDVAEACPT-FLRFEggVpYVTTK---TD 573
 QY 302 SPQARNHPLHVQVKNLDG-----TQYD-----PTD----- 327

Db 574 SDRVLAQDMSLAAKQMSNTFLAGLAQYYTQSGTINLHFMFTGDTDAKARYWVAYAPPG 633
Qy 328 -----DIPAVLGAIDFKGTVFGVASORDVSG 353
Db 634 MEPPKTPAAAHCIHAEDWTGLNSKFTSIP-YLSAADYATASGVAETTNVQG 686

RESULT 10

US-08-583-562B-8
; Sequence 8, Application US/08583562B
; Patent No. 5922570
; GENERAL INFORMATION:
; APPLICANT: Staunton, Donald
; APPLICANT: Harris, Edith
; TITLE OF INVENTION: Cytoplasmic Modulators of Integrin
; TITLE OF INVENTION: Binding
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583.562B
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/33033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2647 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-583-562B-8

Query Match 4.1%; Score 119; DB 2; Length 2647;
Best Local Similarity 21.0%; Pred. No. 0.074;
Matches 131; Conservative 64; Mismatches 219; Indels 209; Gaps 31;
Qy 48 IDPWIREN-----FVQAPQ-GEFTVSPR-----NSPGEMLLNLGLPGLNPYL--S 90
Db 1696 VDDVVENEDGTDFITYTAPQPGKYICVRFGGHEVPNPPQVITALAGDQPSVQPLRSQ 1755
Qy 91 HLSRMVNGYACGMQVQVVLNAGNATAGKIIIPAAVPPHPPVENISAAQITMCPHIVDVQ 150
Db 1756 QLAPQYTVAGGQQTWA-----PERPLGVNG-----LDVTS 1787
Qy 151 LEPLLPLPDIRNRRFFHYNQENTPRML-----VAMLYTPLRANGSE- 192
Db 1788 LRPDLVIP-----FTIKKGITGEVRMPGSKVAQPTITDNKDGTVTVRYAPSEAGLHEM 1842
Qy 193 -----DVFTVSCRVLTRPADPFEFTFLVPPTVSSKTKPFTLPILTIGEL 236
Db 1843 DIRYDNMHIPGSLQFYVDVYNCGRHTAYGPG-----LTHGVNKKPATFTVNTKDAGEG 1896
Qy 237 SNS---RPPAIDMLYTDPN-----SIWVQ-----PQNGRCITLDG 269
Db 1897 GLSLAIEGFSKAEISCTDNQDGTCSVSLPVLPGDYSILVKNBQHVPGSPFTARTVGTDD 1956

Qy 270 TLQ-----GTTQLVPTQICAFRGTLLISQTARAADSTDSP---QRARN----- 308
Db 1957 SMRSHLKVGSAAIDIPINISSETDLSLTATVVPSPGREEPCLLKLRLNRHGVHSFVPKXET 2016
Qy 309 --HPLHVQKNLQGTQYDPTDDIPAVLGAIDFKGTVFGVASORDVSGQ-QFQGHVATRAH 365
Db 2017 GEHLVHVVK-KN---GQHVASSPIPVISOSE-----IGDASRVVSGGGLHEGHTFEPA- 2066
Qy 366 EAHITDTPKYPKGLGILIKSGSD--DFNTNQPIRETTPVGMGDNMNRQWBLPDSGRLT 423
Db 2067 EFLIDTRDAGY---GLSLSTIEGSPKVDINTED-----LEDGTCRVTYCPTPEGNYI 2115
Qy 424 LNMNLAPAVSPSP-----GE-----RILPFRSIVPSAGGYSGY-IDCLIPQEWVQH- 471
Db 2116 INIKFADQHVSGSPFSVKVTGEGRVKESITRRRAPSIVANVSGHCDLSLKPIEISIQDWT 2175
Qy 472 YQEAAPS-----QSAVALRVYVNPDTGNIFEAKLHREGPLTVANGCNPPIVVP 520
Db 2176 AQVTSQSGKTHEAIEVEGENTHYCIRFPVPAEMGHTTVSVYKYGQHV----- 2221
Qy 521 PNGYFRFEAMGNQFYTLAPMGSG 543
Db 2222 PGSPQF-----TVGPLGEG 2236

RESULT 11

US-08-779-113-8
; Sequence 8, Application US/08779113
; Patent No. 5948891
; GENERAL INFORMATION:
; APPLICANT: Staunton, Donald E.
; APPLICANT: Harris, Edith S.
; TITLE OF INVENTION: Cytoplasmic Modulators of Integrin
; TITLE OF INVENTION: Binding
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/779.113
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Greta E. No. 5948891and
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/33773
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2647 amino acids
; TYPE: amino acid
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
US-08-779-113-8

Query Match 4.1%; Score 119; DB 2; Length 2647;
Best Local Similarity 21.0%; Pred. No. 0.074;
Matches 131; Conservative 64; Mismatches 219; Indels 209; Gaps 31;
Qy 48 IDPWIREN-----FVQAPQ-GEFTVSPR-----NSPGEMLLNLGLPGLNPYL--S 90
Db 1696 VDDVVENEDGTDFITYTAPQPGKYICVRFGGHEVPNPPQVITALAGDQPSVQPLRSQ 1755

us-09-926-799-10.ra1

Wed Jun 2 09:13:24 2004

```

; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1079 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-929-967-8

Query Match      3.8%; Score 109.5; DB 2; Length 1079;
Best Local Similarity 19.3%; Pred. No. 0.15;
Matches 116; Conservative 74; Mismatches 229; Indels 181; Gaps 25;

QY 91 HLSRMVNYAGMQVQVVLGNAGTAGKIIFAAVPPHPFVVENISAAQITMCPHVIVDVQR 150
DB 1756 QLAFOYTVAGGQQTWA-----PERPLGVNG-----LDVTS 1787
QY 151 LEPVLLPLDIRNFFHYNQENTPRML-----VAMLYTPLRANSGE- 192
DB 1798 LRPFDLVP-----PTIKKEITGEVRMPGSKVAQPTITDKDGTVTVRYPAPSEAGLHEM 1842
QY 193 -----DVFTVSCRVLIRPADPFEFTFLVPPTVESKTKPFTPLTLTGL 236
DB 1843 DIRYDNMHPGSPLOFYDYVYNGCHVTAYGPG-----LTHGVVKNKATFTVNTKADAGEG 1896
QY 237 SNS-----RPPAADMLYTDPNE-----SIVVQ-----PQNGRCRTLDG 269
DB 1897 GLSLAIEGSKAEISCTDNQDGTCSVYLPVLPGDYSILVKYNEQHVPGSPFTARVTGDD 1956
QY 270 TLQ-----GTTQLVPTQICAFRGTLSQIOTARAADSTDP---QRA- 308
DB 1957 SRMGMHLKVGSAAIPINISDTLSLTATVVPSPGSEPCLLKRLRNGHVIGISFVPKET 2016
QY 309 --HPLHVQVKNLDGTQYDPTDIPAVLGAIDFKGTVFGVASQDVSGQ-OBQGHYATRAH 365
DB 2017 GSHLVHVK-KN---GOHVASSIPVVISQSE-----IGDASRVVSGQGLHEGHTFEDA- 2066
QY 366 EAHIDTDPKYAPKLTILIKSGSD--DFNTNQPIRFTFVGMGNMNRQWELPDYSGRLT 423
DB 2067 EFLIDTRDAGYG--GLSLSIEGSKVDINTED-----LEDGTCRVTCYCTEPGNYI 2115
QY 424 LNMNLAPAVSPSP-----GE-----RILFFSIVPSAGGYGSGY-IDLCLIPQEWVQHF- 471
DB 2116 INIKFADQHVPGSPFSKVYTGREGVKESITRRRRAPSVANVGHCDLSLKPEISIQDWT 2175
QY 472 YQEAAPS-----QSAVALRYVNPDTGRNIFEAKLHREGFLTVANGCNPIVVP 520
DB 2176 AQVTSFSGKTHAEAEIVGEGNHTYCIKRVPAEMGTHTVSKYKGQHV----- 2221
QY 521 PNGYFRFEAWGNQFYTLAPMGSG 543
DB 2222 PGSPFQF-----TVGPLLEG 2236

RESULT 12
US-08-929-967-8
; Sequence 8, Application US/08929967
; Patent No. 5891637
; GENERAL INFORMATION:
; APPLICANT: Ruppert, Siegfried J.W.
; TITLE OF INVENTION: Construction of Full-Length cDNA Libraries
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/929,967
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1035R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881

```

```

; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1079 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-929-967-8

Query Match      3.8%; Score 109.5; DB 2; Length 1079;
Best Local Similarity 19.3%; Pred. No. 0.15;
Matches 116; Conservative 74; Mismatches 229; Indels 181; Gaps 25;

QY 24 DEVMALPEVVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSPRNSPGEMLNLELGP 83
DB 339 DQKVAQVEIKQALLTAPALGLPLTKPF--ELFVDEKQY-----AGVITQKLG 387
QY 84 ELNPYLSHSRMYNGYAGG---MQVQVVLGNAGTAGKIIFAAVPPHPFVVENISAAQIT 139
DB 388 WRRP-VAYLSKCLDPVAGWPPCLRWAAIAVLTKDAGL-----TMGQPLVI 434
QY 140 MCPHVIVDVQRLEPVLPLPD--IRN-RFFHYNQENTPRMLVAMLYTPLRANSGEDVPT 196
DB 435 LAPHAV-----EALVKQPPDRWLSNARMTHYQ-----ALLDTRVQGFVVAL 478
QY 197 VSCRVLTRPADPFEFTFLVPPTVESKTKPFTPLTLTGLSRSRPPAIDMLYTDPNESI 256
DB 479 NPATLLPLPEEGLOHNCILDILABAHGTRP-----DLTDQPLPDADHTWTYDGS 529
QY 257 VVQPNQRCRTLDGTQGTQTLV-----PTQICAFRGTLI--SQTARAADS----- 303
DB 530 -----QSGQKAGAAVTEVIMAKALPAGTSAQRAELIALTQALKMAEGKLNVTDSR 585
QY 304 QRARNHPLHVQV-----KNLDGTQYDPTDIPAVLGAI----- 336
DB 586 YAFATAHIHGEIYRRRGLLTSEKTEIKNKDEILALILKALFLPKRLSIHCPGHQKHS 645
QY 337 -----DFKGTVFGVASQDVSGQEQGHYATRAEAHIDTDPKYAPKLTILIKSGS 389
DB 646 ARGNRMADQAAKAAITETPDTSLIENSPPYSEHFHYTVTDIKDLTKLGAIVDKTKK 705
QY 390 DFNTNQPI--RFTVPMGDNMNRQWELPDYSGRLT-----LNMNLAPAVSPSP- 436
DB 706 YWVYQKPVMPDQFT-----FELDLHLQHLTHLSFSKMKALLERSHSFYMLNR 754
QY 437 -----PGERI-----LPFRSIVPSAGGYGSGYIDC 461
DB 755 DRLKNITETCKACAQVNASKSAVKQGRVGRHPRPGTHWEIDFTEIKP--GLYGYKLLV 812
QY 462 LIP--QEWVQHFYQEAAPSQSAVALRYVNPDTGRNIFEAKLHREGFLTVANGCNPIV 519
DB 813 FIDTFSGWI-----EAFPTKKTAKV-----VTKKLLEIFRFGMPQVLGTDNGPAFV 861

RESULT 13
US-09-309-572-13
; Sequence 13, Application US/09309572
; Patent No. 6440730
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral hybrid vectors pseudotyped with LCMV
; FILE REFERENCE: P50489
; CURRENT APPLICATION NUMBER: US/09/309,572
; CURRENT FILING DATE: 1999-05-11
; EARLIER APPLICATION NUMBER: DE 198 56 463
; EARLIER FILING DATE: 1998-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1737
; TYPE: PRT
; ORGANISM: Moloney murine leukemia virus
; FEATURE:
; OTHER INFORMATION: gag-pol protein
; US-09-309-572-13

```

Query Match 3.8%; Score 109.5; DB 4; Length 1737;
 Best Local Similarity 19.3%; Pred. No. 0.33;
 Matches 116; Conservative 74; Mismatches 229; Indels 181; Gaps 25;

Qy 24 DEVNLEPVVGCASIAAPVVGQNIIDPWIRENFVQAQGEFTVSPRSPGEMLLNLELGP 83
 Db 997 DQKAYEIQKALLTAPALGLPDLTKPF--ELFVDEKQY-----AKGVLTQKLG 1045
 Qy 84 ELNPSYLSHLSMYNGYAGG---MQVQVVLGNAFTAGKIIFAAVPPHPPVENISAAQIT 139
 Db 1046 WRRP-VAYLSKKLDPVAGWPPCLRMVAAIAVLTKDAGK-----TWGQPLVI 1092
 Qy 140 MCPHIVDVROLEPVLLPLPD--IRN-RFFHYNOENTPRMELVAMLYTPLRANGSEDEVFT 196
 Db 1093 LAPHAV-----EALVKOPPPDRWLSNARMTHYQ-----ALLDTRVQFGPVVAL 1136
 Qy 197 VSCRVLTRPAPDFEFTFLVPPTVESKTPFLPILTLGELSNSRFPAAIDMLYTDPNESI 256
 Db 1137 NPATLLPPEGLQHNCILDLAEAGTRP-----DLTDQPLPDADHTWYTDGSSLL 1187
 Qy 257 VVQPNQGRCTLDGTLQGTTLV-----PTQICAFRGTLI--SQTARAADS-----TDSP 303
 Db 1188 ---QEGORKAGAAVTTETEVIAKALPAGTSAQRAELIALTQALKWAEKGLNVYTD 1243
 Qy 304 QARNHPLHVQV-----KNLDGTQYDPTDIPAVLGAI-----336
 Db 1244 YAFATAIHGEIYRRRGLLTSEGKEIKNKDBILALLKALFLPKRLSIHCPGHQGHSAE 1303
 Qy 337 ---DFKTVGVASQRDVSCQEQGHVATRAHEAHIDTTPKYAPKLGTLIIKSGS 389
 Db 1304 ARGNRMAQAAKAAITETPTDSTLLIENSPPYTSSEHPHYTVTDIKOLTKLGAIYDKTK 1363
 Qy 390 DDFNTNOPI---RFTPVGMGNRNQWELPDYSGRLT-----LNMNLAPAVSPSF--- 436
 Db 1364 YWVYQGVPMVDQPT-----FELLDFLHLQTLHLSFKMKALLERSHSPYMLNR 1412
 Qy 437 -----PGERI-----LFFRSIVPSAGYSGYIDC 461
 Db 1413 DRTLKNITETCKACAQVNASKSAVKQGTVRVGRHPRGTHWEIDFTEIKP--GLYGYKYL 1470
 Qy 462 LIP--QBWVQHFYQEAAPSQSAVALVRYVNPDTGRNIFPEAKLHREGFLTVANCNNPIV 519
 Db 1471 FIDTFSGWI-----EAPFTKETAKEV-----VTKKLEIEIFPRFGMPQVLGTNDGPAFV 1519

RESULT 14

US-09-718-096-13
 ; Sequence 13, Application US/09718096
 ; Patent No. 6589763
 ; GENERAL INFORMATION:
 ; APPLICANT: Von Laer, Meike-Dorothee
 ; TITLE OF INVENTION: RETROVIRAL HYBRID VECTORS PSEUDOTYPED WITH LCMV
 ; FILE REFERENCE: 35-195
 ; CURRENT APPLICATION NUMBER: US/09/718,096
 ; CURRENT FILING DATE: 2000-11-22
 ; PRIOR APPLICATION NUMBER: DE 19856463.5
 ; PRIOR FILING DATE: 1998-11-26
 ; PRIOR APPLICATION NUMBER: EP 99250415.9
 ; PRIOR FILING DATE: 1999-11-25
 ; PRIOR APPLICATION NUMBER: US 09/309,572
 ; PRIOR FILING DATE: 1999-05-11
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 13
 ; LENGTH: 1737
 ; TYPE: PRT
 ; ORGANISM: Moloney murine leukemia virus
 ; FEATURE:
 ; OTHER INFORMATION: gag-pol protein
 US-09-718-096-13

Query Match

3.8%; Score 109.5; DB 4; Length 1737;

Best Local Similarity 19.3%; Pred. No. 0.33;
 Matches 116; Conservative 74; Mismatches 229; Indels 181; Gaps 25;

Qy 24 DEVNLEPVVGCASIAAPVVGQNIIDPWIRENFVQAQGEFTVSPRSPGEMLLNLELGP 83
 Db 997 DQKAYEIQKALLTAPALGLPDLTKPF--ELFVDEKQY-----AKGVLTQKLG 1045
 Qy 84 ELNPSYLSHLSMYNGYAGG---MQVQVVLGNAFTAGKIIFAAVPPHPPVENISAAQIT 139
 Db 1046 WRRP-VAYLSKKLDPVAGWPPCLRMVAAIAVLTKDAGK-----TWGQPLVI 1092
 Qy 140 MCPHIVDVROLEPVLLPLPD--IRN-RFFHYNOENTPRMELVAMLYTPLRANGSEDEVFT 196
 Db 1093 LAPHAV-----EALVKOPPPDRWLSNARMTHYQ-----ALLDTRVQFGPVVAL 1136
 Qy 197 VSCRVLTRPAPDFEFTFLVPPTVESKTPFLPILTLGELSNSRFPAAIDMLYTDPNESI 256
 Db 1137 NPATLLPPEGLQHNCILDLAEAGTRP-----DLTDQPLPDADHTWYTDGSSLL 1187
 Qy 257 VVQPNQGRCTLDGTLQGTTLV-----PTQICAFRGTLI--SQTARAADS-----TDSP 303
 Db 1188 ---QEGORKAGAAVTTETEVIAKALPAGTSAQRAELIALTQALKWAEKGLNVYTD 1243
 Qy 304 QARNHPLHVQV-----KNLDGTQYDPTDIPAVLGAI-----336
 Db 1244 YAFATAIHGEIYRRRGLLTSEGKEIKNKDBILALLKALFLPKRLSIHCPGHQGHSAE 1303
 Qy 337 ---DFKTVGVASQRDVSCQEQGHVATRAHEAHIDTTPKYAPKLGTLIIKSGS 389
 Db 1304 ARGNRMAQAAKAAITETPTDSTLLIENSPPYTSSEHPHYTVTDIKOLTKLGAIYDKTK 1363
 Qy 390 DDFNTNOPI---RFTPVGMGNRNQWELPDYSGRLT-----LNMNLAPAVSPSF--- 436
 Db 1364 YWVYQGVPMVDQPT-----FELLDFLHLQTLHLSFKMKALLERSHSPYMLNR 1412
 Qy 437 -----PGERI-----LFFRSIVPSAGYSGYIDC 461
 Db 1413 DRTLKNITETCKACAQVNASKSAVKQGTVRVGRHPRGTHWEIDFTEIKP--GLYGYKYL 1470
 Qy 462 LIP--QBWVQHFYQEAAPSQSAVALVRYVNPDTGRNIFPEAKLHREGFLTVANCNNPIV 519
 Db 1471 FIDTFSGWI-----EAPFTKETAKEV-----VTKKLEIEIFPRFGMPQVLGTNDGPAFV 1519

RESULT 15

US-09-252-991A-28530
 ; Sequence 28530, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 28530
 ; LENGTH: 335
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-28530

Query Match 3.5%; Score 102.5; DB 4; Length 335;

Best Local Similarity 21.5%; Pred. No. 0.11;
Matches 64; Conservative 37; Mismatches 121; Indels 75; Gaps 12;

Qy 222 KTKPFTLPLTLGELSNSRFPAAIDMLYTDPNESIVVQPNQGRCTLDGTLQGTTLVPTQ 281
 Db 28 QNEPARTFYSQSATARPQPEVEAEFRITPLEAPLGAEVRG-----LDARRPLAPEQ 80

us-09-926-799-10.rai

Wed Jun 2 09:13:24 2004

QY	282	ICAFRGTLISQTARAADSTDSQARAHPLHVQVKNLDGTQY-----DPTDDIP	330
Db	81	VLALKQAL-----REHHILVFRQQHLDDEQYLRFATLFGSVFQPPADIP	124
QY	331	AVLGAIDFK-CTVFGVASQORDVSCQOQGHYATRAHEAHIDTTDPKYAPKLGITLIKSGS	389
Db	125	VLSSGGDGKVPD1VKVAN---TGDGELGNFALPAHIDHQTVPVSSGSELYALEVPSSG	180
QY	390	DDFNINQPIRFTPVGMGDNWRQWELPDYSGRLTNNNLAPAVSPSPGGERILFFRSIVP	449
Db	181	GE-----TRFT-----NLARAYESLDEATREID-----GLRLINYPFIR	216
QY	450	-SAGGYGSGYIDCLIPQEWVQHFQEAAPSQ-SAVALVRYVNPDTGRNIFEAKLHRE	504
Db	217	LREGGYGGGFATYRTP-----DIEPIQGSHEPLVR-THPESGRVLFLSAHE	263

Search completed: June 1, 2004, 13:58:09
Job time : 16.104 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 35.1594 Seconds
(without alignments)
4368.312 Million cell updates/sec

Title: US-09-926-799-10
Perfect score: 2906
Sequence: 1 MKMASNDAPNSDGAANLVP.....GNQPYTLAPMGSGGRRRAQ 550

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1141	39.3	530	US-10-314-739-3	Sequence 3, Appli
2	280.5	9.7	547	US-10-670-695-4	Sequence 4, Appli
3	280.5	9.7	668	US-10-209-507-4	Sequence 4, Appli
4	280.5	9.7	671	US-10-670-695-2	Sequence 2, Appli
5	272.5	9.4	669	US-10-209-507-2	Sequence 2, Appli
6	124.5	4.3	1147	US-10-327-481A-38	Sequence 38, Appli
7	121	4.2	1344	US-09-738-626-6888	Sequence 6888, Ap
8	119	4.1	2609	US-10-043-487-583	Sequence 383, App
9	119	4.1	2647	US-10-231-956A-87	Sequence 87, Appli
10	111.5	3.8	1199	US-10-677-558-2	Sequence 2, Appli
11	104.5	3.6	944	US-10-174-677-101	Sequence 101, App
12	104	3.6	368	US-10-156-761-7936	Sequence 7936, Ap
13	103.5	3.6	394	US-10-104-047-3560	Sequence 3560, Ap
14	103	3.5	3317	US-10-080-334-200	Sequence 200, App
15	102	3.5	1167	US-10-282-122A-62618	Sequence 62618, A

16	102	3.5	1184	12	US-10-282-122A-64572	Sequence 64572, A
17	101	3.5	563	12	US-10-425-114-59009	Sequence 59009, A
18	101	3.5	572	12	US-10-425-114-46359	Sequence 46359, A
19	101	3.5	723	12	US-10-425-114-38441	Sequence 38441, A
20	101	3.5	815	12	US-10-425-114-39399	Sequence 39399, A
21	101	3.5	815	12	US-10-425-114-62649	Sequence 62649, A
22	100.5	3.5	410	12	US-10-282-122A-48886	Sequence 48886, A
23	100.5	3.5	1307	15	US-10-104-047-2438	Sequence 2438, Ap
24	100.5	3.5	6620	15	US-10-080-334-290	Sequence 290, App
25	100	3.4	1794	10	US-09-965-738-299	Sequence 299, App
26	100	3.4	1799	10	US-09-965-738-149	Sequence 149, App
27	100	3.4	1821	10	US-09-965-738-82	Sequence 82, Appli
28	100	3.4	2234	12	US-10-612-090-20	Sequence 20, Appli
29	100	3.4	5877	14	US-10-142-515-11	Sequence 11, Appli
30	100	3.4	5935	14	US-10-243-243A-8	Sequence 8, Appli
31	100	3.4	11721	10	US-09-965-738-162	Sequence 162, App
32	99.5	3.4	534	12	US-10-282-122A-64860	Sequence 64860, A
33	99.5	3.4	1257	9	US-09-738-626-4750	Sequence 4750, Ap
34	99.5	3.4	2435	12	US-10-282-122A-47453	Sequence 47453, A
35	99	3.4	836	14	US-10-272-459-40	Sequence 40, Appli
36	99	3.4	980	14	US-10-272-459-41	Sequence 41, Appli
37	99	3.4	1382	14	US-10-010-160-18	Sequence 18, Appli
38	99	3.4	2227	9	US-09-929-955-12	Sequence 12, Appli
39	99	3.4	2227	13	US-10-104-966-12	Sequence 12, Appli
40	99	3.4	2227	13	US-10-135-988-2	Sequence 2, Appli
41	99	3.4	2227	13	US-10-135-988-4	Sequence 4, Appli
42	99	3.4	2227	13	US-10-135-988-6	Sequence 6, Appli
43	99	3.4	2227	16	US-10-719-619-12	Sequence 12, Appli
44	97.5	3.4	713	15	US-10-369-493-708	Sequence 708, App
45	97.5	3.4	2365	14	US-10-156-761-14890	Sequence 14890, A

ALIGNMENTS

RESULT 1
US-10-314-739-3
; Sequence 3, Application US/10314739
; Publication No. US20030129588A1
; GENERAL INFORMATION:
; APPLICANT: Bates, Mary K
; Jiang, Xi
; Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; Characterize No. US20030129588A1walk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Ave., N.W.
; CITY: Washington, D.C.
; STATE: <Unknown>
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/10/314,739
; APPLICATION NUMBER: US/10/314,739
; FILING DATE: 09-Dec-2002
; CLASSIFICATION: <Unknown>
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,049
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311.023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX: <Unknown>

```

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-314-739-3

Query Match          39.3%; Score 1141; DB 14; Length 530;
Best Local Similarity 43.4%; Pred. No. 2.1e-103;
Matches 249; Conservative 78; Mismatches 169; Indels 78; Gaps 12;

QY 1 MKMASNDAPSNDGAA---NLVPEAN-DEVMALEPVVGASIAAPVVGQNIIDPWIRENF 56
DB 1 MMWASKDATSSVDGASGAGQLVPEVNASDPLANDPVAGSSATAVAGQVNPIDWIINF 60

QY 57 VQAPQGEFTVSPRNSFGEMLLNLELGPENLYSLSRMYNGYAGGMQVQVVLGNAGFTA 116
DB 61 VQAPQGEFTISPNTGDLVFLSLGPHLNPFLHLSQWYNGWGVNMRVIRLGNAGFTA 120

QY 117 GKIIFAAVPHFPVENISAAQITMCPHVIQVLEPVLPLPDIRNRRFFHYNQENTPRM 176
DB 121 GKIIIVCIPPGGSHNLTAQATLPHVIVADVTLDPIEVPLEDRVNLFNHNDNRNQTM 180

QY 177 RLVMALYTPLRANSGB-DVFTVSCRVLTRPAPDFEFTFLVPPTVSKTKPFTLPILTGE 235
DB 181 RLVMALYTPLRANSGB-DVFTVSCRVLTRPAPDFEFTFLVPPTVSKTKPFTLPILTGE 240

QY 236 LNSRFPAAIDMLYTPNESIVVQPNQRCITLDGTLQGTQVLTQICAFRGTLLISQATAR 295
DB 241 LNSRPLPISMGISPDNVQSVQFQNGRCTLGRLVGTPTVSLSHVAKIRGT----- 293

QY 296 AADSTDSQPARNHLVQVKNLDGTDYDTPDDIPAVLGAIDFKG-----TVFGVASQ 348
DB 294 -----SNGTVINTELDTGTFPHFEG-PAPIGPDLGGCDWHINMTQFGHSSQ 340

QY 349 RDVSGQOEGHYATRAHEAHDITDPKYAPKLGTILIKS-GSDFNTNQPIRTPVGMGD 407
DB 341 T-----QYDVTDTPTDFVPHLGSIQANGIGSGNY-----VGW 372

QY 408 NNWTFQ-----WELPDYSGRLTLNNMLAPAVSPFPGERILFFRSIVPSAGGYG 455
DB 373 LSWITSPSPHPSGSDVWLKIPNYGSSITEATHLAPSVYPPGFEVLVFFMSKMPGCAYN 432

QY 456 SGYIDCLIPQWVQHFOEAAPOSQSAVALVRYNPDTRGRNIFEAHLHREGFLTVANGC-- 513
DB 433 ---LPCLLPQEVISHLASEQAPTVEAALLHYVDPTGRNLGFEKAYPDGFLTCVPGNGAS 489

QY 514 NNPIVPPNGYFRFAWNGQYTLAPMGSGQGR 547
DB 490 SGPOQLPINGVVFVSWSRFYQLKPVGTASSAR 523

RESULT 2
US-10-670-695-4
; Sequence 4, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4

```

```

; LENGTH: 547
; TYPE: PRT
; ORGANISM: Feline calicivirus
; US-10-670-695-4

Query Match          9.7%; Score 280.5; DB 12; Length 547;
Best Local Similarity 24.1%; Pred. No. 2.6e-18;
Matches 130; Conservative 74; Mismatches 207; Indels 129; Gaps 24;

QY 11 SNDGAANLVPEANDEV--MALEPVVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSP 68
DB 1 ADDGSIT-APEQGTWVGVIAPESQAMSTAAOMATGKSVDSSEW--EAFSP-HTSVNWST 56

QY 69 RNSPGMLNLLELGPENLYSLSRMYNGYAGGMQVQVVLGNAGFTAAGKIIFAAVPHFP 128
DB 57 SETQKILFKQSGLPLNLYLHAKLYVWAGSGSIEVRFSGSGVFGKLAALVVPVPGV 116

QY 129 -PVENISAAQITMCPHVIQVLEPVLPLPDIRNRRFFHYNQENTPRMRLVAMLYTLP- 186
DB 117 DPVQSTMLQY---PHVLEFARQVEFVIFCLPDLRSTLYHL-MSDITDTSLVMVYNDLI 172

QY 187 -----RANSGEDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPFTLPILTGEISNRF 241
DB 173 NPVANDANSGCCIVTVE---TKPGDFEFHLLKPEG-----SMLTHGSIPLSLI 218

QY 242 PAA-----IDMLYTPNESIVVQPNQRCITLDGTLQGTQVLTQICAFRGTLLISQATAR 297
DB 219 PKTSSLWIGNRYNSDITDFVIRP-----FVQANRHF 250

QY 298 D-----STDSQPARNHLVQVKNLDGTDYDTPDDIPAVLGAIDFKGTVFGVASQRDVSGQ 354
DB 251 DFNQETAGSTPRPRPISVTITQONG-----AKLG-----IGVATDYIVPG- 291

QY 355 QEQGHYATRAHEAHDITDPKYAPKLGTILIKS-GSDFNTNQPIRTPVGMGDNNRWQ 413
DB 292 -----IPDGWPDITIPGELIPAGDYAITNGTGNDITATGATDYADIIKNTNFR-- 340

QY 414 ELPDYSGRLLTNMLAPAVSPFPGERILFFRSIVPSA---GGVSGYIDC-LIPOEWV- 468
DB 341 -----GMVYCSLQRAW-GDKKISNTAFTTATLDGNNKINFCNTIDQSKI 388

QY 469 ---QHFOEAAPOSQSAVALVRY-----VNPDTGRNIFEAHLHREGFLTVANGCNPI 517
DB 389 VFQDNHVKKAKTSDDTLALLGYTGIGEAIGSDRDRWRISTLPETG----ARGGNHPI 444

RESULT 3
US-10-209-507-4
; Sequence 4, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
; US-10-209-507-4
; Query Match          9.7%; Score 280.5; DB 14; Length 668;

```


Wed Jun 2 09:13:24 2004

```
RESULT 6
US-10-327-481A-38
; Sequence 38, Application US/10327481A
; Publication No. US20040001864A1
; GENERAL INFORMATION:
; APPLICANT: King, Andrew M.O.
; APPLICANT: Burdman, Alison J.
; APPLICANT: Audonnet, Jean-Christophe F.
; APPLICANT: Lombard, Michel F.A.
; TITLE OF INVENTION: Vaccine against Foot-and-Mouth Disease
; FILE REFERENCE: 454313-3178
; CURRENT APPLICATION NUMBER: US/10/327,481A
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: PCT/FR01/02042
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: FR 0008437
; PRIOR FILING DATE: 2000-06-29
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patent in ver. 2.1
; SEQ ID NO 38
; LENGTH: 1147
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
; US-10-327-481A-38

Query Match      4.3%; Score 124.5; DB 15; Length 1147;
Best Local Similarity 19.3%; Pred. No. 0.023;
Matches 67; Conservative 41; Mismatches 158; Indels 81; Gaps 6;

Qy 70 NSPGEMLNLELGPENLPYLSHLSRMVYAGQVQVVLGNAFTAGKIIIPAAVPPHP 129
Db 158 DKPGYLTLELPTDHHGVGHLVDSYAYMRNGWDEVSAGVQFNGGCLLVAMPEWKA 217
Qy 130 VENISAAQITWCPHVIVDVRLQLEVLPLPDIRNRPYHYNQENTPRMRLVAMLYTTLRAN 189
Db 218 FDTREKQLTLFPHQFI-----SPRNTMTAHTIVYLGW 251
Qy 190 SGEVFTVSCRVLTRPAPDFEFLVPPTVESKTKPFTLPILTLGELSNSRFPAAIDMLY 249
Db 252 NRYDQY-----KXHKPMTLVVMVLSPLTVSNTAAPQIKVY 286
Qy 250 TDPNESIVVQPONGRCTLGDTGTTQLVPTQICAFRGTLISQTAARADSTDSQPARNH 309
Db 287 ANIAPTVV-----HVAGELPSKEGIFPVACADGYGGLVTTDPKTD----- 327
Qy 310 PLHVQVKNLDGTOY-----DPTDDIPAVLGAIDFKGTGFGVVASQDVSQGOEGHYA 361
Db 328 PVYGVKVPPTKNTYVGRFTNLDVAEACPTFLRFDGKPYVVVTRADDTRLAKFDVSLAA 387
Qy 362 TRAHEAHIDTTPDKYAPKLGTI---LIKSGDDDFNTNQPIRFTPVGM 405
Db 388 KHSNTYLSGIAQYTYQSGTINLHFMFTGTSDSKARYWVAYIPPGV 434

RESULT 7
US-09-738-626-6888
; Sequence 6888, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
```

```
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent in ver. 3.0
; SEQ ID NO 6888
; LENGTH: 1344
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-6888

Query Match      4.2%; Score 121; DB 9; Length 1344;
Best Local Similarity 23.1%; Pred. No. 0.065;
Matches 109; Conservative 52; Mismatches 166; Indels 144; Gaps 29;

Qy 23 NDEVMALFVVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSPRNSPGEMLNLELG 82
Db 188 NEAVWAAP--GDVLIQP--GKSLVF--NIK---NGPNDEATAADFNA--EYGTNLEAG 235
Qy 83 PELNPYLSHLSRMVYAGQMVOQ-----VVLAGNAFTAG-KLIIFAAVPPHPVENISAA 136
Db 236 KDLVEISS--GGMANGTARGMQIQTNTGHIVNRGFYNNMAGSDVKANEGHFAVDESLL 293
Qy 137 QITWCPHVIVDVRLQLEVLPLPDIRNRPYHYNQENTPRMRLVAMLYTTLR---ANSGED 193
Db 294 KQTL-----VSGGAP-----TPGTVYTSQIPNPLSAVIADSSVP 327
Qy 194 VFTVSCRVLTRPAPDFEFLVPPTVESKTKPFTLPIL-TLGELS-----NSR 240
Db 328 LITDNTATSNPAPEFTFAENITDDVQVITA--TLHVTSSAGEAATTINITEDDGSFNWA 385
Qy 241 FPA-----IDMLYTDNESIVVQPONGRCTLGDTGTTQLVPTQICAFRGTLIS 291
Db 386 LPAADLTGKSWFEYTVTATDGFNSVTTEPV--RTVDG-----ANTDPLRLNEENQWVS 438
Qy 292 OTARAADSTDSQPARNHPLHVQVKNLDGTOYDPTDDIPAVLGAIDFKGTGFGV-ASQD 350
Db 439 GTTVDIGASD---VFGDKLELLI-----DDAPAVTNSLSAAPTFAMEVTQTD 483
Qy 351 V-----SQQ-----EQHYATRAHEAHIDTTPKY-----APKL 380
Db 484 VFFRNGILLAGGELRIFDQGTY---ANTETISTPVFLYHINEDGTLTVSVVAGTKAAPEI 540
Qy 381 GTILIKSGDDDFNTNQPIR-----FTPVGMGDNNRQW-ELPDYSGR 422
Db 541 D----LNENDDFOIRN-LRLILPDGRTLTTPAGISDSN--AWLNMGDSAGKL 585

RESULT 8
US-10-043-487-383
; Sequence 383, Application US/10043487
; Publication No. US20030055220A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, LEGRAIN
; TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides
; TITLE OF INVENTION: mammalian polypeptides
; FILE REFERENCE: B4778A
; CURRENT APPLICATION NUMBER: US/10/043,487
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/261,130
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 561
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 383
; LENGTH: 2609
; TYPE: PRT
; ORGANISM: Shigella Flexneri
```

US-10-043-487-383

Query Match 4.1%; Score 119; DB 14; Length 2609;
Best Local Similarity 21.0%; Pred. No. 0.3;
Matches 131; Conservative 64; Mismatches 219; Indels 209; Gaps 31;

QY 48 IDPWIREN-----FVQAPQ-GEFTVSPR-----NSPGMLNLELGPBLNPLYL--S 90
DB 1658 VDQVVENEDGTDFIYTAPOGKVICVRFGEHVPNPFOVTTALAGDQPSVQPPLSRQ 1717
QY 91 HLSRMYNGYAGGMQVQVVLGNAGNFTAGKIIFAAVPPHPFVENISAAQITMCPHVIVDVQR 150
DB 1718 QLAPOYTAQGGQQTWA-----PERPLVGNG-----LDVTS 1749
QY 151 LEPVLLPLDIRNRFHYNQENTPMRL-----VAMLYTFLRANSGB- 192
DB 1750 LRPFDLVIP-----FTIKKGEITGEVRMPGSKVAQPTITDNKDGTVTVRYAPSEAGLHEM 1804
QY 193 -----DVFTVSCRVLTRPADPFEFTFLVPPVTSKTKPFTLPILTLGEL 236
DB 1805 DIRYDNMHIPGSPLOFYDYVNCGHVTAYGPG-----LTHGVNKNPATFTVNTKDAGEG 1858
QY 237 SNS-----RPPAIDMLYTDPE-----SIWVQ-----PQNGRCTLDG 269
DB 1859 GLSLAIEGSKAEISCTDNQDGTCSVSYLPVLPGDYSILVKYNEQHVPGSPPTARVTGDD 1918
QY 270 TLQ-----GTTQLVPTQICAPRGHTLISQARAADSTDS-----QRARN----- 308
DB 1919 SNRSHLKVGSAAIPINISSETDLSLTATVVPSPGREBPCLLKRLRNGHVGISFVPKET 1978
QY 309 --HPLHVQVKNLDGTQYDPTDDIPAVLGAIDPKGTGVFVASORDVSGQ-QEKGHYATRAH 365
DB 1979 GEHLVHVK-KN---GQHVASSPFPVVISQSE-----IGDASRVVSGQGLHGHTEPEA- 2028
QY 366 EAHIDTTPKAPKLTGTLIKSGSD--DENTNQPIRFTPVGMDNNRWQWELPDYSGRLT 423
DB 2029 EFIIDTRDAGY---GLSLSIEGSKVDINTED-----LEDGTCRVYCTEPGNYI 2077
QY 424 LNMNLAPAVSPSPF-----GE-----RILFFRSIVPSAGGYSGY-IDCLIPQEWVQHF- 471
DB 2078 INIKPADQHVPGSPSVKVTGEGRVKESITRRRRAPSVANVGHCHDLSLKPEISIQDWT 2137
QY 472 VQEAAPS-----QSAVALVRYNPDTRGNI FEAKLHREGFLTVANGCNPPIVVP 520
DB 2138 AQVTSFGSKTHAEIVEGENHTYCIIRFVPAEMGTHTVSVKYGQHV-----2183
QY 521 PNGYFRFEAWGNQFYTLAPMGSG 543
DB 2184 PGSPFQF-----TVGPLGEG 2198

RESULT 9

US-10-231-956A-87

; Sequence 87, Application US/10231956A
; Publication No. US2004005323A1
; GENERAL INFORMATION:
; APPLICANT: Lorens, James B.
; APPLICANT: Xu, Weiduan
; APPLICANT: Bogenberger, Jakob
; APPLICANT: Holland, Sacha
; APPLICANT: Rigel Pharmaceuticals, Incorporated
; TITLE OF INVENTION: Modulators of Angiogenesis
; FILE REFERENCE: 021044-004100US
; CURRENT APPLICATION NUMBER: US/10/231,956A
; CURRENT FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 87
; LENGTH: 2647
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-231-956A-87

Query Match

4.1%; Score 119; DB 12; Length 2647;
Best Local Similarity 21.0%; Pred. No. 0.3;
Matches 131; Conservative 64; Mismatches 219; Indels 209; Gaps 31;

QY 48 IDPWIREN-----FVQAPQ-GEFTVSPR-----NSPGMLNLELGPBLNPLYL--S 90
DB 1696 VDQVVENEDGTDFIYTAPOGKVICVRFGEHVPNPFOVTTALAGDQPSVQPPLSRQ 1755
QY 91 HLSRMYNGYAGGMQVQVVLGNAGNFTAGKIIFAAVPPHPFVENISAAQITMCPHVIVDVQR 150
DB 1756 QLAPOYTAQGGQQTWA-----PERPLVGNG-----LDVTS 1787
QY 151 LEPVLLPLDIRNRFHYNQENTPMRL-----VAMLYTFLRANSGB- 192
DB 1788 LRPFDLVIP-----FTIKKGEITGEVRMPGSKVAQPTITDNKDGTVTVRYAPSEAGLHEM 1842
QY 193 -----DVFTVSCRVLTRPADPFEFTFLVPPVTSKTKPFTLPILTLGEL 236
DB 1843 DIRYDNMHIPGSPLOFYDYVNCGHVTAYGPG-----LTHGVNKNPATFTVNTKDAGEG 1896
QY 237 SNS-----RPPAIDMLYTDPE-----SIWVQ-----PQNGRCTLDG 269
DB 1897 GLSLAIEGSKAEISCTDNQDGTCSVSYLPVLPGDYSILVKYNEQHVPGSPPTARVTGDD 1956
QY 270 TLQ-----GTTQLVPTQICAPRGHTLISQARAADSTDS-----QRARN----- 308
DB 1957 SNRSHLKVGSAAIPINISSETDLSLTATVVPSPGREBPCLLKRLRNGHVGISFVPKET 2016
QY 309 --HPLHVQVKNLDGTQYDPTDDIPAVLGAIDPKGTGVFVASORDVSGQ-QEKGHYATRAH 365
DB 2017 GEHLVHVK-KN---GQHVASSPFPVVISQSE-----IGDASRVVSGQGLHGHTEPEA- 2066
QY 366 EAHIDTTPKAPKLTGTLIKSGSD--DENTNQPIRFTPVGMDNNRWQWELPDYSGRLT 423
DB 2067 EFIIDTRDAGY---GLSLSIEGSKVDINTED-----LEDGTCRVYCTEPGNYI 2115
QY 424 LNMNLAPAVSPSPF-----GE-----RILFFRSIVPSAGGYSGY-IDCLIPQEWVQHF- 471
DB 2116 INIKPADQHVPGSPSVKVTGEGRVKESITRRRRAPSVANVGHCHDLSLKPEISIQDWT 2175
QY 472 VQEAAPS-----QSAVALVRYNPDTRGNI FEAKLHREGFLTVANGCNPPIVVP 520
DB 2176 AQVTSFGSKTHAEIVEGENHTYCIIRFVPAEMGTHTVSVKYGQHV-----2221
QY 521 PNGYFRFEAWGNQFYTLAPMGSG 543
DB 2222 PGSPFQF-----TVGPLGEG 2236

RESULT 10

US-10-677-558-2

; Sequence 2, Application US/10677558
; Publication No. US20040096972A1
; GENERAL INFORMATION:
; APPLICANT: AUDIT, Muriel
; APPLICANT: COSSET, Francois-Loic
; TITLE OF INVENTION: CHIMERIC PLASMID COMPRISING A REPLICATIVE RETROVIRAL GENOME AND
; FILE REFERENCE: 1759.135
; CURRENT APPLICATION NUMBER: US/10/677,558
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: PCT/FR02/03934
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: FR 0114976
; PRIOR FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1199
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: product of pol gene of pAM plasmid
US-10-677-558-2

Db	30	QLRYSVVESEPGTLVGNVAQDLGLKMTDLLSRLQLGSENGRYFSLSLMSGALAVNQK	89
Qy	105	VQVVLGNAFTAGKIIFAAVPPHPPVENISAAQTMCPHVIVDVROLEPVLPLPDIRNR	164
Db	90	IDRESLCGASTSCLLPVQVWTEH-PLELI	132
Qy	165	PFHYNQENTPMRLVAMLYTPLRANSGBDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTK	224
Db	133	FATPEREMRISSAASGARFPLDSAQDPDVGNTVSVFYL-SPNSHFSNLVKTLKDG--K	189
Qy	225	PFTLITLIGELSNRFPAAIDMLYDNE	273
Db	190	PF--PELVLEQQLDREAQARHQLVLTAVDGGTTPARSGLTISVILVLDINDNAP	244
Qy	274	TTQLVTOICAFRGTLISQTABAADSTSPQARNHPL	325
Db	245	SVLRVGIPEPNAFTIGTLRL--LNATD-PDRGTNGQLDYSFGDHTSEAVRNLFQ--LDP	297
Qy	326	TDDIPAVLGAIDFKGTGVFGVVASQDVSGQEQGHYATRAHEAHIDTTPKYPAPKLTILI	385
Db	298	SSGAIHVLGPIDFE	336
Qy	386	XSGSDDFNTNQ	432
Db	337	QVDVGDVNDNAPEVLLASLANPVLESTPVGTVVGLFNVDRD--SGR--NGEVSLDI	389
Qy	433	SPSPGGERILFFRSIVPSAGGYG	464
Db	390	SPDLFPQ	442
Qy	465	QEWQHFYQEAAPQSQSAVALRVYVNPDTGRTNIPEAKLHREGFLTVAN	511
Db	443	NISVDNDNAPRFNQQLYTAVILENRRPPGSLCTVAASDPDTGN	493
Qy	512	CNNPIVVPPNGY	524
Db	494	VGNQVQAGAPASSF	506
RESULT 12			
US-10-156-761-7936			
; Sequence 7936, Application US/10156761			
; Publication No. US20030119018A1			
; GENERAL INFORMATION:			
; APPLICANT: OMURA, SATOSHI			
; APPLICANT: IKEDA, HARUO			
; APPLICANT: ISHIKAWA, JUN			
; APPLICANT: HORIKAWA, HIROSHI			
; APPLICANT: SHIBA, TADAYOSHI			
; APPLICANT: SAKAKI, YOSHIYUKI			
; APPLICANT: HATTORI, MASAHIRA			
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES			
; FILE REFERENCE: 249-262			
; CURRENT APPLICATION NUMBER: US/10/156,761			
; CURRENT FILING DATE: 2002-05-29			
; PRIOR APPLICATION NUMBER: JP 2001-204089			
; PRIOR FILING DATE: 2001-05-30			
; PRIOR APPLICATION NUMBER: JP 2001-272697			
; PRIOR FILING DATE: 2001-08-02			
; NUMBER OF SEQ ID NOS: 15109			
; SEQ ID NO 7936			
; LENGTH: 368			
; TYPE: PRT			
; ORGANISM: Streptomyces avermitilis			
US-10-156-761-7936			
Query Match 3.6%; Score 104; DB 14; Length 368;			
Best Local Similarity 23.2%; Pred. No. 0.38; Indels 112; Gaps 18;			
Matches 82; Conservative 39; Mismatches 120; Indels 112; Gaps 18;			
Qy	16	ANLYPEANDEWMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAPGEGFTVSPRNSPGEM	75

Db	24	DEVMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAPGEGFTVSPRNSPGEMLINLELGP	83
Qy	459	DOQKAYQETKQALLTAPALGLPDLTKPF--ELFVDEKQY	507
Db	84	ELNPYLSHLSRMVYAGS---MQVVVLGNAFTAGKIIFAAVPPHPPVENISAAQIT	139
Qy	508	WRPP--VAYLSKKLDPAAGWPPCLRWAAIAVLTKDAGL	554
Db	140	MCPHVIVDVROLEPVLPLPD--IN-REFHYNQENTPMRLVAMLYTPLRANSGBDVFT	196
Qy	555	LAPHAV-----EALVKQPDRLNLSNARWTHQ	594
Db	197	VSCRVLTRPAPDFEFTFLVPPTVESKTPFTLPI	251
Qy	595	---VVALNPA-----TL--PLPEEGLQHDCLDILAEAGHTRSDLTDOPLPADHTWYTD	644
Db	252	PNBSIVVQONGRCTLDGTGTLQV---	299
Qy	645	SSSFL-----QEQKAKAGAAVTETEVIVARALPAGTSAQRAELIALTQALKWAECKLVN	700
Db	300	TDSPQARNHPLHVQV---KNLDGTQYDPTDDIPAVLGAI	336
Qy	701	YDSRVAFATAHGEIYRRGLTSEGKEIKNKDEILALLKALFLPKRLSIHCPGHQK	760
Db	337	---DFKGTGVFVASQDVSGQEQEQ---	384
Qy	761	GNSAARGNRMAQAAAREVATRETPTGSTLLIENSTPYTHEHPTVTTDKLTKLGATY	820
Db	385	IKSGSDDFNTNQI---RFTPVGMDNNRWQWELPDYSGRLT--LNNMLAPAV---SPS---	435
Qy	821	DSAKKYVYQKVPWPDQFT-----FELDLFLHQLTHLSFSKTKALLERSPSPY	869
Db	436	-----PFGERI-----LFFRSIVPSAGGYGS	456
Qy	870	YMLNRDRLKNIETCKACAQVNASKSAVKQGTVRGHRFGTHWEIDFTEVKP--GLYGY	927
Db	457	GYIDCLIP--QEWQHFYQEAAPQSQSAVALRVYVNPDTGRTNIPEAKLHREGFLTVANCGN	514
Qy	928	KYLLVFDVTFSGWI-----EAPFKKETAKV-----VTKKLLBEEIPFRGMPQVLGTDN	976
Db	515	NPVV 519	
Qy	977	GPAFV 981	
RESULT 11			
US-10-174-677-101			
; Sequence 101, Application US/10174677			
; Publication No. US20030190704A1			
; GENERAL INFORMATION:			
; APPLICANT: Xie, Ting			
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ANCHORING STEM CELLS IN A MICROENVIR			
; FILE REFERENCE: 40716(IP-012)			
; CURRENT APPLICATION NUMBER: US/10/174,677			
; CURRENT FILING DATE: 2002-06-19			
; NUMBER OF SEQ ID NOS: 117			
; SOFTWARE: Patentin version 3.1			
; SEQ ID NO 101			
; LENGTH: 944			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-174-677-101			
Query Match 3.6%; Score 104.5; DB 14; Length 944;			
Best Local Similarity 21.5%; Pred. No. 1.6; Indels 165; Gaps 22;			
Matches 119; Conservative 56; Mismatches 213; Indels 165; Gaps 22;			
Qy	61	QGEFTVSPRNSPGEMLIN-----LELGPFLNPLYLSHLSRMVYAGGQM 104	

Db 56 AELLPOG-----ALDPPARAQIVLTVGRSTVLMGVTRHEVARSAE-----PQQA 102
Qy 76 LLNLGLPELNPYLHLGRM-----YNGYAGGMQVQVVLGNAGFTAGKIIFAAVP 125
Db 103 KLTL-TGED-----VSRMDVDFSGVKYVGLPPEERVAVIVA--KYTAGLVPAIVP 152
Qy 126 PHF-----PVENISAQITMCPHVIDVROLEP-----VLLPLDIRRFFHYNGEN 172
Db 153 SVFMVAPPLQAVPHSHQDTLAY-----VRQLAAQAGYQFFVAPGLPGV--NIAHMBPPN 206
Qy 173 TPRLMLVAMLYTPLRANGSE-----DVFTVSCRVLTRPADPFEFTFLVP-PTVESKTKP 225
Db 207 KGLQDOELLVNCDAASNVESLSFSDGFSATQYVVLQNPKTKFPVPVPVDPVSPPLSP 266
Qy 226 FTLPLITLGLSNSRFPAAIDMLYTDPNESIVVQPNCRCTLDGLOGTTQLVPTQICAF 285
Db 267 M-----GRRRLPLK-----VQP-----IHGLAKLTPQAAGV 294
Qy 286 RGLTISQARAADSTD-----SPQARNH-PLHVQVKNLDGTQY 323
Db 295 -----ALARAADSFEVSGGSLDVLRYGQPLRARPPVPHGAGAAYDGTYY 341

RESULT 13

US-10-104-047-3560
; Sequence 3560, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3560
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3560

Query Match 3.6%; Score 103.5; DB 15; Length 394;
Best Local Similarity 22.7%; Pred. No. 0.48;
Matches 49; Conservative 32; Mismatches 66; Indels 69; Gaps 9;
Qy 129 PVENISAQITMCPHVIDVROLEPVLLPLP-----DIRRFFHYNGENTPRMLVAML 182
Db 73 PNPVSCATPSLLSLVLKCPCLSLIPLPGRKCTYGIKCRFFPBRPSCPQ-RSVA-- 129
Qy 183 YTPLRANGSDVFTVSCRVLTRPADPFEFTFLVPPTVESKTKFPFLPILTLGELNSRFP 242
Db 130 -DELRA-----LLSPRPSKDK-----NGRRP 153
Qy 243 AAIDMLYTDPNESIVVQPNCRCTLDGLOGTTQLVPTQICAFRGTLLISQARAADSTD 302
Db 154 -----SPSSQSSLLTSEQCSLDGKLG-AQASP-----GSRQEGLTQTVAPSGRSLA 201
Qy 303 PQARNHPLHVQVKNLDGTQDPTDDIPAVLGAIDF 338
Db 202 PSGG-----SGSSFGPTDLPQLTDLSPY 225

RESULT 14

US-10-080-334-200
; Sequence 200, Application US/10080334
; Publication No. US20040002584A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shinkets, Richard A
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
Query Match 3.5%; Score 103; DB 15; Length 3317;
Best Local Similarity 19.2%; Pred. No. 17;
Matches 136; Conservative 73; Mismatches 242; Indels 256; Gaps 34;
Qy 14 GAANLVPEANDEVMALEFVVGASIAAPVVGQNIIDPWIRENFVQAP---QGEFTVSPRN 70
Db 1703 GRYTLIVTATDQC-----FILSHRLTSTTTVLNVND--INDNVPTFPRDYGEPFDVT-BG 1755
Qy 71 SPGEMLLNLGLPELNPYLHLGRMSTVLMGVTRHEVARSAE-----PQQA 130

1756 QP-----GPRVWTFPLAH--DRDSCPNQVEYSVV---DGDPLGEFVISPVEGLRV 1801
131 -----ENISAAQITMCPH-----VIVVURQLEPVLLPDPDIRNFFH 167
1802 RKDVELDRETIAFYNLITICARDRGVPLLSSTMLVGIRVLDINDNDPVLNLP----- 1853
168 YN--QENTPRMVLY-----AMLYTPLRANSQDVF-----TVSCRVLTRP-- 205
1854 MNITISENSPVSSFFVAHVLASDADGCGNALLTFNITAGNRREAFINATTGIIVTNPLD 1913
206 ---APDEFETFLV---PPTVESKTPFTPLITLGELSNRPFAAIDMLYTDPNESIVVQ 259
1914 RERIEFRLTVSKONPENPRIARKDFDLVSLAD-ENDNHP-----LFTEGTYQAEVM 1967
260 PONGRCLTDLTQGG-----TTOLVPTQICAF-----RGLTISQ 292
1968 ENSPACTPLTVLNGPILALDADEVVAVVYQLLGTHSDLFVIDNSTGVTVRSVILDR 2027
293 TABAADSTQSPQRARNHPLHVKNLDGTQY-----DPTDDIPAV----- 332
2028 EA-----FSPFLELLLAEDVGQLNGTAYLFTITLDDNDNWPFTSPPAYTVHLENC 2081
333 ---LGAIDFKGVFGVASORDVSGQEQG--HYAT---RAHEAHD----- 370
2082 PGFSVLQITATDSDGLNGELVYRIEAGAQDRFLIHPVTGVRVGNATIDREEQESYRLT 2141
371 --TDPKYAPKLTILIKSGSDDFNTNQP-----IRFTPVGMGD 407
2142 VVATDRGTVPLSGTATVITLIDINDSRPEFNPIQTVSVLESTEGTIVANVTAILDL 2201
408 NNRQWEL-----PQYSGRLTLMNLA----- 429
2202 NPKLEYHILSVAKDDTDLRVDPQEDAFVAVNTGTVSVIKSNRLVATYETVLTSLVIN 2261
430 ---PAVSPSPGGR-----ILFPRSIVPSAGGSGYIDCLIPQEWQHFYQAAAPSQA 481
2262 ASDLPERSVSPNAKLTWNILDVNDNTPOKPGIITY-----YTERVLEGATPGTTL 2313
482 VALVRYNVPDTRGRNIFEAKLHREGFLTVANGCNPPIVVPNGYFRFE 528
2314 IA-VNAVDPDKGLN-----GLITVTLDD-----LIPP-GYVQLE 2345

RESULT 15
US-10-282-122A-62618
; Sequence 62618, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62618
; LENGTH: 1167
; TYPE: PRT
; ORGANISM: Mycobacterium bovis
US-10-282-122A-62618

Query Match 3.5%; Score 102; DB 12; Length 1167;
Best Local Similarity 19.3%; Pred. No. 3.9;
Matches 105; Conservative 73; Mismatches 211; Indels 154; Gaps 26;

QY 19 VPEANDEWMALEPVVGASTAAAPVVGQONIIDPWIRENFVQAPQGEF-----TVSPR 69
DB 310 VPDGHDQASWL-----RHEVDAGLRRRFPAGPDGGRERAAYEIDVICSK 354
QY 70 NSPEMLLNLELGPENLPYLHLRVYAGVGMQVVLVAGNAFTAGKIIFAAPVPHFP 129
DB 355 GFPSYFLIVADL-----ISVARSAGIRVG-PGRGSAAGSLVAYA----- 392
QY 130 VENISAAQITMCPHVIYDVRLQLEFVLLPLDPI-----RNRFFHYNQENTPRMLV-- 179
DB 393 ---LGITDIDIPHGLLPERFLNPERTSPMDIDIDFDRRERGENVRYAADKMGHDRVAVQ 449
QY 180 -----AMLYTPLRANSQDVFVTSRVLTRPADPDEFETFLVPPTVESKTKPFTLPI 230
DB 450 ITFGTIKTKAALKDSARIHYQPGFAIDRI-----TKALPPAIMAK-----DPL 495
QY 231 LTLGELSNRFPAAID---MLYTDPNESIVVQPNQRCCTLDGTLOGTQLVPTQICAFRG 287
DB 496 SGITDPSHERYKEAAEVRGLIETDPDVRTIYQTARG-----LEGLRNAGVHACA--- 545
QY 288 TLISOTARAADSTDSFQARNHPLHVQVKNLDGTQYDPTDDIPAVLGAIDFKG-----TVF 343
DB 546 -VIMSSEPLTEAIPLMKRPQDGA-----ITGWDYPACEAIG--LLKMDFLGLRNLTI 596
QY 344 G-----VASORDVSGQEQGHVATRAHEAHDITTDPKYAPKLTILIKSG--SDDFNTNQ 396
DB 597 GDAINDVRANRGIDLDLESVPLDDKATVELLGRGD-----TLGVFQDGGPMRDLRRMQ 651
QY 397 PIRFT-----PVGMDNNRWQWELPDYSGRLTLNMLNLAAPVSPGPFGERILFF 444
DB 652 PTGFEDVAVIALYRPGPMGMAHN-----DYADR-----KNNQAIKPIHP-ELEBPL 699
QY 445 RSIVPSAGVSGYIDCLIPQEWQHFYQAAAPSQSAVA-LVRYNVPDTRGRNIFEAKLHR 503
DB 700 REIL--AETVG-----LIVYQIMRIAKQVASYSLARADILRKAMGKKREVLKEF-- 750
QY 504 EGF 506
DB 751 EGF 753

Search completed: June 1, 2004, 14:04:46
Job time : 37.1594 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:34:56 ; Search time 6.6443 Seconds
(without alignments)
4310.252 Million cell updates/sec

Title: US-09-926-799-10
Perfect score: 2906
Sequence: 1 MKMASNDAPSDGAAANLVP.....GNQFYTLAPMGSGGRRRAQ 550

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1830.5	63.0	539	1 COAT LORDV	P54635 lordedale v
2	1089.5	37.5	546	1 COAT SOUV3	Q04542 southampton
3	293	10.1	702	1 COAT SMSV1	P36284 san miguel
4	290	10.0	2344	1 POLN RHDV	P27410 rabbit hemo
5	287.5	9.9	703	1 COAT SMSV4	P36285 san miguel
6	283	9.7	668	1 COAT FCVF4	P27405 feline cali
7	280.5	9.7	671	1 COAT FCVF9	P27406 feline cali
8	270.5	9.3	668	1 COAT FCVC6	P27404 feline cali
9	236.5	8.1	2208	1 POLN MANCV	Q69014 manchester
10	129.5	4.5	2332	1 POLG FMDV0	P03305 f genome po
11	125.5	4.3	2194	1 POLG HE701	P32537 h genome po
12	124.5	4.3	2333	1 POLG FMDV1	P03306 f genome po
13	119	4.1	2647	1 FLNA HUMAN	P21333 homo sapien
14	118.5	4.1	2193	1 POLG CX16T	Q9qf31 c genome po
15	118	4.1	2336	1 POLG FMDV2	P49303 f genome po
16	117.5	4.0	3375	1 UNS2 CABEL	Q06561 caenorhabdi
17	117	4.0	4289	1 TENX HUMAN	P22105 homo sapien
18	116	4.0	2193	1 POLG CX16G	Q65900 c genome po
19	115	4.0	1034	1 ADD DROME	P54362 drosophila
20	114.5	3.9	788	1 BCBS XANAC	P58933 xanthomonas
21	114	3.9	2194	1 POLG EC30B	Q9w78 e genome po
22	114	3.9	2647	1 FLNA MOUSE	Q8btm8 mus musculus
23	113	3.9	855	1 POLG HRV3	Q82081 human rhino
24	111.5	3.8	2175	1 POLG BOVEV	P12915 b genome po
25	111.5	3.8	2206	1 POLG POL32	P06209 poliovirus
26	110.5	3.8	1204	1 POL MLVFS	P26810 friend muri
27	110.5	3.8	1204	1 POL MLVFF	P26809 friend muri
28	110.5	3.8	2332	1 POLG FMDVA	P03308 f genome po
29	109.5	3.8	1199	1 POL MLVMO	P03355 moloney mur
30	108.5	3.7	798	1 PBPA NEICI	Q86088 neisseria c
31	108.5	3.7	1204	1 POL MLVFP	P26808 friend muri
32	107	3.7	1011	1 POLG FMDVT	P15072 foot-and-mo
33	107	3.7	2201	1 POLG CXA9	P21404 c genome po

RESULT 1									
COAT LORDV	COAT LORDV	STANDARD;	PRT;	539	AA.				
AC	P54635;								
DT	01-OCT-1996	(Rel. 34, Created)							
DT	01-OCT-1996	(Rel. 34, Last sequence update)							
DT	01-NOV-1997	(Rel. 35, Last annotation update)							
DE	Coat protein (Capsid protein).								
OS	Lordedale virus (Human enteric calicivirus).								
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;								
OX	Norovirus.								
OX	NCBI_TaxID=82658;								
RN	[1]	SEQUENCE FROM N.A.							
RP	MEDLINE=96005060; PubMed=7561776;								
RA	Dingle K.E., Lambden P.R., Caul E.O., Clarke I.N.;								
RT	"Human enteric Caliciviridae: the complete genome sequence and								
RT	expression of virus-like particles from a genetic group II small								
RT	round structured virus."								
RL	Genbank accession: F662395-2355(4995)								
CC	- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.								
CC	-----								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
CC	the European Bioinformatics Institute. There are no restrictions on its								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial								
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/								
CC	or send an email to license@sib-sib.ch).								
CC	-----								
DR	EMBL; X86557; CAA60255.1;								
DR	InterPro; IPR004005; Calici_coat.								
DR	InterPro; IPR008975; Viral_Cap_coat.								
DR	Pfam; PF00915; Calici_coat; 1.								
KW	Coat protein; Glycoprotein.								
FT	CARBOHYD 172 172								
FT	CARBOHYD 307 307								
FT	CARBOHYD 406 406								
FT	CARBOHYD 531 531								
SQ	SEQUENCE 539 AA; 58775 MW; 5E5C6E7F2C5PD21 CRC64;								
Query Match 63.0%; Score 1830.5; DB 1; Length 539;									
Best Local Similarity 62.1%; Pred.No. 2.3e-130;									
Matches 346; Conservative 83; Mismatches 101; Indels 27; Gaps 6;									
QY	1	MKMASNDAPSDGAAANLVP	EAANLVP	EAANLVP	EAANLVP	EAANLVP	EAANLVP	EAANLVP	EAANLVP
DB	1	MKMASNDAPSDGAAANLVP	EAANLVP	EAANLVP	EAANLVP	EAANLVP	EAANLVP	EAANLVP	EAANLVP
QY	61	QGEFTVSPRNSPGEMLLNLELPELNPYL	SHLSRM	YAGM	QVQV	VL	AGNA	FTAG	KI 120
DB	61	QGEFTVSPRNSPGEMLLNLELPELNPYL	SHLSRM	YAGM	QVQV	VL	AGNA	FTAG	KI 120
QY	121	FAAVPHPHFVENISAQITMCPHVIVD	VRQLEP	VL	PL	PD	INR	FFHY	QNTFPRMLVA 180
DB	121	FAAVPHPHFVENISAQITMCPHVIVD	VRQLEP	VL	PL	PD	INR	FFHY	QNTFPRMLVA 180

Matches	250; Conservative	75; Mismatches	173; Indels	89; Gaps
QY	1	MMASNDAAASNDGAA---	NLVPEAND-EVMALFPPVVGASIAAPVVGQOMLIDPWIRENF	56
Db	1	MMASNDAAASNDGAA---	NLVPEAND-EVMALFPPVVGASIAAPVVGQOMLIDPWIRENF	60
QY	57	VOAPOGEFTVSPRNSPGMMLNLGLPELNPYLSHLSRMVYAGGQVQVVLGNAPTA	116	
Db	61	VOSPOGEFTISPNNTPGDILLDLQGLPHLNPFLSHLSQMYNGVGNRVRILLAGNAPSA	120	
QY	117	GKIIFAAAPPHEFVENISAAQITMCQPHVIVVQRLPVLPLPDIRNRFHYNENTPRM	176	
Db	121	GKIIICCVCPGPTSSSLTIAQATLFPFHVIAVTRTEIEMPLEDVRNVLVYTN-DNQPTM	179	
QY	177	RLVAMLYTLPLRANS---	EDVFTVSCRVLTRPAPDPEFTFLVPPVTSKTPFTLPIITL	233
Db	180	RLVCMLYTLPLRTGGSGNSDSFVAGRVLTAPSSDFSLFLVFPPTIEQKTRAFVFNIP	239	
QY	234	GELSNRFPAAIDMLYTPNPSISVVPQNGRCTLDTGLTQGLTQVPTQICAFRGLTISQT	293	
Db	240	QTLNSRFPFSLIQGMILLSPDASQVQFQNGRCLIDQLLGTTPATSGQLFRVRGK-ING	298	
QY	294	ARAADSTDSQARHPLHVQVKNLDGTQYDPTDDIPAVLGAIDF-----	KG 340	
Db	299	ARTLNLT-----	VDGKPFMAFDS-PAPVGPDPFGKCDWMHRISKTPNN	341
QY	341	TVFG-----	VASORDVSGQOEGHYATRAHEAHIDTTPKYAPKGLTILKSGSDDFN--	393
Db	342	TGSGDPMRSVSQVNTVQ---	FVPHLGSIOF---DEVNHP	376
QY	394	-----	TNQPTRFTPVGMGDNWRQWELPDVSGRLTLNNMLAPVSPGGERILF	443
Db	377	TGDYIGTIEWISQP--STPPGTIN---	LWEIPDYSSLSQAANLAPPVFPFGGEALV	431
QY	444	FRSIVPSAGYSGS-GYIDCLIPQEWVQHYQEAAPSQSAVALVRYVNPDPTRGRIFEAKH	502	
Db	432	FVSAPFGPNNSAPNDVPCLLPQEVITHFVSEQAPTMDAALLHYVDPDNRNLGEPKLY	491	
QY	503	REGFLTVA--NCGNNPIVPPNGYFRFEAWGNQVFTLAPMGSGQGR	547	
Db	492	PGGYLTCPVNGVAGVQPLNGVFLFVSVWVSRYQLKPVGTASTAR	538	
RESULT 3				
COAT SMSV1	STANDARD;	PRT;	702 AA.	
ID	COAT SMSV1			
AC	P36284;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	Coat protein (Capsid protein).			
OS	San Miguel sea lion virus (serotype 1) (SMSV 1).			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;			
OC	Vesivirus.			
OX	NCBI_TaxID=36406;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92410750; PubMed=1529644;			
RA	Neill J.D.;			
RT	"Nucleotide sequence of the capsid protein gene of two serotypes of			
RT	San Miguel sea lion virus: identification of conserved and non-			
RL	conserved amino acid sequences among calicivirus capsid proteins.";			
RL	Virus Rep. 24:211-222(1992).			
CC	-/- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; L07418; AAA92984.1; --			
DR	PIR; B37491; B37491.			
DR	InterPro; IPR004005; Calici_coat.			
DR	InterPro; IPR008975; Viral_cap_coat.			
DR	Pfam; PF00915; Calici_coat; 1.			
KW	Coat protein; Glycoprotein.			
FT	CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT	CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT	CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT	CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).			
SQ	SEQUENCE 546 AA; 58774 MW; C82B2A85AD4B05EA CRC64;			
Query Match	37.5%;	Score	1089.5;	DB 1; Length 546;
Best Local Similarity	42.6%;	Pred.	No.	1.6e-74;


```

DR EMBL; M87481; AAA16217.1; -.
DR FIR; A48562; A48562.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 493 493 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 702 AA; 77850 MW; B6E5A58523DDE3D7 CRC64;

Query Match 10.1%; Score 293; DB 1; Length 702;
Best Local Similarity 25.0%; Pred. No. 2.3e-14;
Matches 117; Conservative 61; Mismatches 150; Indels 140; Gaps 21;

QY 8 AAPSNDGAANLVPEANDEVMALPEVVGAS-----IAAPVVGQNIIDPWIRENFVQAPQGE 63
DB 151 AESDGGGADIVTEEGTVVQGPVPAQSAITTLAASTG-----KTVDCG 196

QY 64 FTV-----SPRNSPEMLLNLELGPENLPYLHSLRMYNGVAGGMQVQVLAGNA 113
DB 197 WTTFFSYHTAVNWSSTTEAQKILFSRALSPENLPYLHSLRMYNGVAGGMQVQVLAGNA 256

QY 114 FTAGKILFAVPPHPPVENISAAQITMCPHVIVDVRQLEPVLPLDINRPFHYNQENT 173
DB 257 VFGGKLAALIVPP--GIEPVESPTMLQYPHVLFDAQTPEVIFTIPDIRKTLHY-SMDDT 313

QY 174 PMRLVAMLYTPI--RANGSEDVFTVSCVLRPADEPDEFTFLVPPTVSKTKPFTLPIL 231
DB 314 DITRLVIMYNELINDEYSEPKSCSIVETRPSSDFTFSLKLP--GSLLKHGSIPTD 371

QY 232 TLGELS---NSRFPAAIDMLYTDPNESIVQPO-----NGRCTLDGTGTTQLVPTQIC 283
DB 372 LIPRNSRHMGNRWWTID-----GFVQPRVFSQNRHFDFTTGWMS--TPYVI- 420

QY 284 AFRGTLISQTARAADSTDSQRAARNHPLHVQKNLD-GTQYD-----P 325
DB 421 -----PIEVTEKLDKRGQYFKVTDTEKSLVPGLPDGWP 454

QY 326 TDDIPAVL---GAIDFKGTGVGASQDVSGQEQGHVATRAHAHIOTDTPKYAPKLG 381
DB 455 DTTIPTAMTASNGYDITVAEYRITN-----NGTHFKGFYI-----MG 492

QY 382 TILIK-SGSDDF-NTNQPIR--FTPVGMGDNWNRQELPDYSGLRLTN 425
DB 493 NLTKVKYKSGDNLGETQTSRTLPASVG-----NYKXDQNTIN 528

RESULT 4
POLN RHDV STANDARD; PRT; 2344 AA.
AC P27410;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Non-structural polyprotein (Contains: RNA-directed RNA polymerase
DE (EC 2.7.7.48); Thiol protease P3C (EC 3.4.22.-); Helicase (2C like
DE protein); Coat protein)
OS Rabbit hemorrhagic disease virus (RHDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Lagovirus.
OC NCBI_TaxID=11976;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91361557; PubMed=1840711;
RA Meyers G., Wirblich C., Thiel H.-J.;
RT "Rabbit hemorrhagic disease virus -- molecular cloning and nucleotide
RT sequencing of a calicivirus genome.";
RL Virology 184:664-676 (1991).
CC -!- FUNCTION: THE P3C CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE
CC FOR CLEAVAGE AT CERTAIN Q/G SITES IN THE POLYPROTEIN.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +

```

```

{RNA} (N).
-!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC EMBL; M67473; AAA47285.1; -.
DR FIR; A41039; RNRWRH.
DR PDB; 1KHV; 16-JAN-02.
DR PDB; 1KHW; 16-JAN-02.
DR MEROPS; C24.001; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR000317; Peptidase_C24.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_P5vir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR Pfam; PF03510; Peptidase_C24; 1.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00916; 2CENDOPTASE.
DR PRINTS; PR00918; CALICIVIRUSNS.
DR SMART; SM00382; AAA; 1.
KW Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;
KW Thiol protease; Helicase; ATP-binding; Coat protein; 3D-structure.
FT CHAIN 2508 2707 HELICASE (P2C-LIKE).
FT CHAIN 1109 1251 PROTEASE P3C.
FT CHAIN 7149 ?1625 RNA-DIRECTED RNA POLYMERASE.
FT CHAIN ? 2344 COAT PROTEIN.
FT NP BIND 522 529 ATP (POTENTIAL).
FT ACT_SITE 1212 1212 PROTEASE (POTENTIAL).
FT ACT_SITE 1227 1227 PROTEASE (POTENTIAL).
SQ SEQUENCE 2344 AA; 257066 MW; 1454C248F81E9212 CRC64;

Query Match 10.0%; Score 290; DB 1; Length 2344;
Best Local Similarity 30.3%; Pred. No. 1.9e-13;
Matches 89; Conservative 49; Mismatches 114; Indels 42; Gaps 12;

QY 8 AAPSNDGAANL-----VPEANDE-----VMALPEVV-----GASIAAPVVG-----QNIID 49
DB 1772 AAPQGEAAGTATTASVPGTTDGMDFGVVATTSVITAENSSASIAATAGIGGPPQVDOOE 1831

QY 50 PWIRENFVQAPQGEFTVSPRNSPGEMLLNLELGPENLPYLHSLRMYNGVAGGMQVQVVL 109
DB 1832 TW-RTNFY--NDVFTWSVADAPGSILYTVQHSPPNNPFTAVLSQMYAGWAGMQRRPIV 1888

QY 110 AGNAFTAGKIIIPAAVPPHPPVENISAAQITMCPHVIVDVRQLEPVLPLDINRPFHYN 169
DB 1889 AGSGVFGGRLVRAVIPP--GIEIGPGLVVRQPHVVIDARSLEPVTITPDLRPNMYHT 1946

QY 170 QENTPRMLVAMLYTPIPLRANGSEDVFTVSCVLRPADEPDEFTFLVPPTVSKTKP- 225
DB 1947 GDPGLVPTLVLSVYNNLINPFQGSATIQVTVETRPSEDFEFVIMRAPSKTVDSISPAG 2006

QY 226 -FTLPILTGLGELNSRFPAAIDMLYTDPNESIVQPOG-----RCTLDGTGTT 275
DB 2007 LUTTPVLT-GVGNDRWNGQI-----VGLQPVPGFSTCNRHWNLNGST 2049

RESULT 5
COAT_SMSV4
ID COAT_SMSV4 STANDARD; PRT; 703 AA.
AC P36285;
DT 01-JUN-1994 (Rel. 29, Created)

```

01-JUN-1994 (Rel. 29, Last sequence update)
 15-DEC-1998 (Rel. 37, Last annotation update)
 Coat protein (Capsid protein).
 San Miguel sea lion virus (serotype 4) (SMV 4).
 Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 Vesivirus.
 NCBI_TaxID=36407;
 [1]
 SEQUENCE FROM N.A.
 MEDLINE=92410750; PubMed=1529644;
 Neill J.D.;
 "Nucleotide sequence of the capsid protein gene of two serotypes of
 San Miguel sea lion virus: identification of conserved and non-
 conserved amino acid sequences among calicivirus capsid proteins.";
 Virus Res. 24:211-222(1992).
 -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 EMBL; M87482; AAA16220.1; -.
 PIR; C48562; C48562; Calici coat.
 InterPro; IPR008975; Viral_cap_coat.
 Pfam; PF00915; Calici_coat; 1.
 Coat protein; Glycoprotein.
 CARBOHYD 89 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 482 482 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 703 AA; 77721 MW; CSDAD8223B261073 CRC64;
 Query Match 9.9%; Score 287.5; DB 1; Length 703;
 Best Local Similarity 24.5%; Pred. No. 5.9e-14;
 Matches 104; Conservative 67; Mismatches 144; Indels 109; Gaps 18;
 8 AAPSDNGAANLVPEANDEVMALEPVVG-----ASTAAPVVGQNIIDPWIRENFVQAPQGE 63
 151 AESDGFSGAEIVTERQGTWVQQQPAPAPATATLATATAGTG-KSVEQEWMTFFSYHTSINW 209
 64 FTVSPRNSGEMLLNLELGPENLPYLHSLSRMYNGYAGGMQVQVVLGNAFTAGKIIFAA 123
 210 STV---ESGKILYSQALNPSINPYLDHIAKLYSTWGGIDVFTVSGSGVFGKLAALL 266
 124 VPPHF--PVENISAAQITMCPHVIVDVRQLEPVLPLDIRNRPFFHYNQENTPRMLVAML 182
 267 VPPGVPIESVSMQLQ---PHVLFDAQVPEVIFTIPDIRKTLFH-SMDETDTTKLVINP 322
 183 YTPLRANSEDVFTVSCRVLTPAPDFETELVPP----- 217
 323 Y-----ENGVENKTCISITVETPSADFTFALLKPPGSLIKGSI1PSDLIPNSAHWMGNR 378
 218 -----TVESKTKPFTLP--ILTLGEL--SNSRFPADIMLY 249
 379 WNSTISGFSVQPRVFNQRHFDSTTTGWSSTPYYPYPIKIKGKVGSKNKFHVID--- 435
 250 TD-----PNESIVQPNQGRCTLDGTLQ-GTTLVPTQICA-FRGTILIS 291
 436 TDKALVREGIPDGWPTTIPDET---KATNGNFSGESYRAGSTIKPNENSTHFKTYIC 492
 292 QTARAADSDSPKARNRHPHVVQVKNLDGTQYDPTDDIPAVLGAIDFKGTVFGVASQRDV 351
 493 GTLSTVETPEENDEQ-----QIKTEAEKSKQTMVVT-----ADFKDTI--VKPQHKI 537
 352 SQQ 355
 538 SPQ 541

RESULT 6
 COAT_FCVF4 STANDARD; PRT; 668 AA.
 ID COAT_FCVF4
 AC P27405;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Coat protein (Capsid protein).
 OS Feline calicivirus (strain Japanese F4) (FCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OK NCBI_TaxID=11980;
 RN [1]
 RP SEQUENCE FROM N.A.
 MEDLINE=91306470; PubMed=1853578;
 RA Tohya Y., Taniguchi Y., Takahashi E., Utagawa E., Takeda N.,
 Miyamura K., Yamazaki S., Mikami T.;
 "Sequence analysis of the 3'-end of feline calicivirus genome.";
 Virology 183:810-814(1991).
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 EMBL; D90357; BAA14371.1; -.
 PIR; B40481; VCMWFC.
 InterPro; IPR004005; Calici coat.
 InterPro; IPR008975; Viral_cap_coat.
 Pfam; PF00915; Calici_coat; 1.
 Coat protein; Glycoprotein.
 CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 668 AA; 73589 MW; 85BBD8C85804E503 CRC64;
 Query Match 9.7%; Score 283; DB 1; Length 668;
 Best Local Similarity 24.2%; Pred. No. 1.2e-13;
 Matches 131; Conservative 78; Mismatches 198; Indels 134; Gaps 27;
 11 SNDGAANLVPEANDEV--NALEPVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSP 68
 125 ADDGSIT-TPEQGTAVGGVIAEPPSAQMSTADMASGKSVDSM--EAFPSF-HTSVNWST 180
 69 RNSPFGEMLLNLELGPENLPYLHSLSRMYNGYAGGMQVQVVLGNAFTAGKIIFAAVPPHF 128
 181 SETQGLKFKQSLGGLPLNPLYLHSLKLYVWAGSIEVRSISGSGVFGKLAIAVPPGV 240
 129 -PVENISAAQITMCPHVIVDVRQLEPVLPLDIRNRPFFHYNQENTPRMLVAMLYTPL- 186
 241 DPVQSTSMQLQ---PHVLFDAQVPEVIFTIPDIRKTLHV-MSDTRDTTSLVIMVNDLI 296
 187 -----RANSGEDVFTVSCRVLTPAPDFETELVPPFTVETSKTKPFTLITLIGELNSRF 241
 297 NPYANDNSSGCIVTVE---TKPGDPFKHLLKPPG-----SVLTHGSI1PSDLI 342
 242 PAALDM-----LYTDPNESIVQPNQGRCTLDGTLQGTTLVPTQICAFRGTILISOTARA 296
 343 PKSSSLMIGNRYWTDITD-FVIRP-----FVQANRH 373
 297 AD---STDSPPQARNRHPHVVQVKNLDGTQYDPTDDIPAVLGAIDFKGTVFGVASQRDV 353
 374 FDFNQETAGWSTPRPRPITITISEKNGSK-----LG-----IGVATDYIIP 415

```

QY 354 QOEGHYATRAHEAHIIDTDPKYPAPKL-----GTLIKSG-SDDPNTNQPRTFTVGMGDDNN 409
DB 416 -----IPDGWPDFT---IADKLIIPAGDYSITTEGNDIKTAQAYDTAAVVKNTTN 462
QY 410 WRQWELPDYSGRLTLNNNLAPAVSPFPGERILFFRSIVPSAGGYG-----SCYID---C 461
DB 463 FR-----GMYTCGSLQRAW-GDKKISNTAFITTAIRDGNEIKPSNTIDMTKL 508
QY 462 LIPQEWQHYQBAAPSQSAVALRVYVNDTGRNIFEAKLHREGFLTV-----ANGGNP 516
DB 509 AVYQD--THVEQSVQSDTLALAGYTG--IGEEAIGSNRDRVVRISVLPEAGARGGNHP 564
QY 517 I 517
DB 565 I 565

RESULT 7
ID COAT_FCVF9 STANDARD; PRT; 671 AA.
AC P27406;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Coat protein (Capsid protein).
GN CYP76.
OS Feline calicivirus (strain F9) (FCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=11981;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=92410623; PubMed=1529544;
RA Carter M.J., Milton I.D., Meanger J., Bennett M., Gaskell R.M.,
RA Turner P.C.;
RT "The complete nucleotide sequence of a feline calicivirus.";
RL Virology 190:443-448(1992).
[2]
SEQUENCE FROM N.A.
RX MEDLINE=92339673; PubMed=1633955;
RA Meanger J., Carter M.J., Gaskell R.M., Turner P.C.;
RT "Cloning and sequence determination of the feline calicivirus strain
RT F9.";
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M86379; AAA79327.1; -.
DR EMBL; Z11536; CAA77636.1; -.
DR PIR; B43382; VCMWF9.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 618 618 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 671 AA; 73441 MW; 33BEE86D8370D5E5 CRC64;

Query Match 9.7%; Score 280.5; DB 1; Length 671;
Best Local Similarity 24.1%; Pred. No. 1.9e-13;
Matches 130; Conservative 74; Mismatches 207; Indels 129; Gaps 24;
QY 11 SNDGAANLPEANDEV--MALEFVVGASTAAPVVGQNIIDPWIRENFVQAPQGEFTVSP 68

```

```

DB 125 ADDGSIT-APEQGTWVGVIAPSPSAQMSTAADMATCKSVDSSE--BAFFSF-HTSVNWST 180
QY 69 RNSPGMLNLNLGLPELNPYLSHLSRMVNGYAGGMQVQVVLGNAGNAFTAGKIIFAAVPPHF 128
DB 181 SETQKILFKQSLGPLNLNPLYLEHLAKLYVANGSIEVRSISGSGVFGKLAIVVPPGV 240
QY 129 -PVENISAAQITCMCHVIVDVRLPEVLLPLDIRNRFFHYHNOENTPRMLVAMLYTPL- 186
DB 241 DPVQSTSMLOQ---PHVLFDAQVEVIFCLPDLRSTLYHL-MSDTRDTTSLVIMVNDLI 296
QY 187 -----RANGSEDVFTVSCRVLTPAPDPFEFTFLVPPTVESKYPFTLPILITIGELNSRF 241
DB 297 NPYANDANSSGCIIVTVE---TKPGDPFKHLLKPG-----SMLTHGSIPLSDLI 342
QY 242 PAA-----IDMLYTDPNESIVVQNGRGTCTLDGTGTTQTLVPTQICAFRGTLLISQTA 297
DB 343 PKTSSLWIGNRYWSDITDFVIRP-----FVQANRHF 374
QY 298 D---STDSPQARNHPLHVQVKNLDGTQYDPTDDIPAVLGAIDPKGTGFGVASQRDVSGQ 354
DB 375 DFNQETAGWSTPRFRPISVTITEQNG-----AKLG-----IGVATDYIVPG- 415
QY 355 QEOGHYATRAHEAHIIDTDPKYPAPKLGTILIKSGS-DDPNTNQPRTFTVGMGDDNNRWQW 413
DB 416 -----IPDGWPDFTTIPGELIPAGDYAITNGTNDITATTGYDTADIIKNNTNFR-- 464
QY 414 ELPDYSGRLTLNNNLAPAVSPFPGERILFFRSIVPSA---GGYSGGYIDC-LIPOEWV- 468
DB 465 -----GMYTCGSLQRAW-GDKKISNTAFITTAIRDGNNKINPCNTIDOSKIV 512
QY 469 ----QHFYQEAAPSQSAVALRVY-----VNPDTGRNIFEAKLHREGFLTVANGGNPFI 517
DB 513 VFQDNHVGKKAQTSDDTLALGYTGIGEOAIGSDRDRVVRISVLTPLPETG----ARGGNHPI 568

RESULT 8
ID COAT_FCVF6 STANDARD; PRT; 668 AA.
AC P27404;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein (Capsid protein).
OS Feline calicivirus (strain CFI/68 FIV) (FCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=11979;
[1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=91374597; PubMed=1716692;
RA Neill J.D., Reardon I.M., Heinrichson R.L.;
RT "Nucleotide sequence and expression of the capsid protein gene of
RT feline calicivirus.";
RL J. Virol. 65:5440-5447(1991).
[2]
SEQUENCE FROM N.A.
RA Neill J.D.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32819; AAA42925.1; -.
DR EMBL; U13992; AAC13993.1; -.
DR PIR; A40507; VCMWFF.

```

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC InterPro; IPR004005; Calici_coat.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00915; Calici_coat; 1.
CC Coar protein; Glycoprotein.
DR CARBOHYD 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
DR CARBOHYD 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 439 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 668 AA; 73550 MW; 9E52312108D441 CRC64;

Query Match 9.3%; Score 270.5; DB 1; Length 668;
Best Local Similarity 31.1%; Pred. No. 1e-12;
Matches 75; Conservative 41; Mismatches 94; Indels 31; Gaps 10;

QY 11 SNDGAANLVPANDEV--MALEPVVGASIAAPVVGQNIIDPWIRENFVQAGQFTVSP 68
DB 125 ADDGSIT-TPEQGTVMGVIAEPNAQMSTAADMATGKSDSEV--EAFSPF-HTSVNNWT 180

QY 69 RNSPGEMLNLELGPENLNPVLSRMVNGYAGGMQVQVVLGNAFTAGKIIFAAVPPHF 128
DB 181 SETQGIKLFQSLGFLNPLTHLAKLYVAMSGVDVRSISGSGVFGKLAIVVPGI 240

QY 129 -PVENISAAQITWCPHVIVDROLEPVLPLPDIRNFRFFHYNOENTPRMLVAMLYPL- 186
DB 241 DFIGQSTSLQY---PHVLFDAQVEPFIPIPLRLSTLXHL-MSDITDTTSLVIMVYNDLI 296

QY 187 -----RANSGEDVFTVSCRVLTPAPDFEFTFLVPPTVESKTPFTLPLTLGELNSRF 241
DB 297 NPYANDNSNGSCIVTVE---TRKGPDPFKHLLKPPG-----SMLTGHSIFSDLI 342

QY 242 P 242
DB 343 P 343

RESULT 9
POLN MANCV STANDARD; PRT; 2208 AA.
AC Q69014;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: RNA-directed RNA polymerase
DE (EC 2.7.7.48); Thiol protease 3C (EC 3.4.22.-); Helicase (2C like
DE protein); Coat protein].
OS Manchester virus (Human enteric calicivirus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Sapovirus.
OX NCBI_TaxID=82659;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95390791; PubMed=7661689;
RA Liu B.L., Clarke I.N., Caul E.O., Lambden P.R.;
RT "Human enteric caliciviruses have a unique genome structure and are
RT distinct from the Norwalk-like viruses."
RL Arch. Virol. 140:1345-1356(1995).
CC -!- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).
CC -!- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR
CC THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC EMBL; X86560; CAA60262.1; -.
CC MEROPS; C24.UPW; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR00317; Peptidase_C24.
DR InterPro; IPR00605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_P3D.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR Pfam; PF03510; Peptidase_C24; 1.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00916; 2CENDOPTASE.
DR PRINTS; PR00918; CALICIVIRUSN.
KW Polypeptide; Transferase; RNA-directed RNA polymerase; Hydrolase;
KW Thiol protease; Helicase; ATP-binding; Coat protein.
FT CHAIN ? ? HELICASE (P2C-Like).
FT CHAIN ? ? PROTEASE (P3C).
FT CHAIN ? ? RNA-DIRECTED RNA POLYMERASE.
FT CHAIN ? 2208 COAT PROTEIN.
FT NP_BIND 408 415 ATP (POTENTIAL).
FT ACT_SITE 1097 1097 PROTEASE (BY SIMILARITY).
FT ACT_SITE 1112 1112 PROTEASE (BY SIMILARITY).
FT SEQUENCE 2208 AA; 242736 MW; 3E299D5BA20E45DC CRC64;

Query Match 8.1%; Score 236.5; DB 1; Length 2208;
Best Local Similarity 22.3%; Pred. No. 1.9e-09;
Matches 135; Conservative 80; Mismatches 264; Indels 127; Gaps 23;

QY 16 ANLVEANDEVMALEP-----VVGASIAAPVVGQNIIDPWIRENFVQAGQ 61
DB 1653 SNPEKQSNPMVDPPTGTTTSHVVVANPEQNGAAQRLAVATGATQSNVPEAIR 1712

QY 62 GEFTV-----SPRNSPGEMLNLELGPENLNPVLSRMVNGYAGGMQVQVVLGNAFT 115
DB 1713 NCFVAFRTFAWDRMPTGTFLGSLSHFNINPYTHLSGMWAGMGSGFEVLSISGSGVF 1772

QY 116 AGKIIFAAVPPHFFVENISAAQITWCPHVIVDROLEPVLPLPDIRNFRFFHYNOENTPR 175
DB 1773 AGRIIASVIPP--GVDPSIRDPGVLPFAVDARITEPVSFMI PDVRAVDVHRMDGABPT 1830

QY 176 MRLVAMLYTP-LRANSGEDVFTVSCRVLTPAPDFEFTFLVPPTVESKTPFTLPLT 232
DB 1831 CSLGFVWYQPLNPFRTTAVSTCWVSVETKPGGDFCLLRPPQQMGENGVSPEGLLPR 1890

QY 233 LGELNSRPPAIDMLYTPDNESIVVQPNQGRCTLDGLTQGTTLQVPTQICAFRGTLSQ 292
DB 1891 LGYSRGNRVGGLVVGMLVAEHKQVNRHFNNSVTFGWSVTPNPMASIVTNOA---HS 1947

QY 293 TARAADSTDSPPQARNHPLHVQVKNLDGTQYDPTDDIPAVLGAIDPKGVFGVASORDYS 352
DB 1948 TSHAWLSICAQ--NKGPLFPPIPN-----HPDSCASTVVGAMD---TSLGGRPSTGVC 1997

QY 353 G-----QOEGHYATRAHEAHIITDPKAPKLT-----TLKSGSDDFNT 394
DB 1998 GPAISFQNGDVYENDTPSVMFATYDPLTS---GTGVALTNSINPALSALVRISNDFDT 2054

QY 395 NQPIRFTVGM-GDNN---WRQWELPDYSGRLLTNLNLAPAVSP----- 434
DB 2055 S-----GFANDKNVVQMSWEM--YTGTNQIRGQVTPMSGTNTVFTTGTANTLVLWQ 2104

QY 435 ----SFFGERILPFRSIVPSAGGSGYIDCLIPQEVVQHFYQEAAPSQSAVALRVNVP 490
DB 2105 ERMLSYDGHQAILYSSQLERTAEYFQNDI-VNIPENSMVAF-----NV 2146

QY 491 DTGRNIFEAKLRHSGFLTVANCGNNPITVPPNGVFRFEANGNQFYTLAPMGS----- 542
DB 2147 ETNSASFOIGRIPDGMVWG--GSGVNVFLEPETF--GIVGVNVPLEPETF-----QVVGILPLSAAALSGPSGN 2199

QY 543 -GQRR 547
 Db 2200 MGRARR 2205

RESULT 10

POLG_FMDVO STANDARD; PRT; 2332 AA.
 AC P03305;
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat
 DE protein VP4; Coat protein VP2; Coat protein VP3; Coat protein VP1;
 DE Core protein P12; Core protein P14; Core protein P14; Genome-linked
 DE protein VP3; Protease (EC 3.4.22.-); RNA-directed RNA polymerase
 DE (EC 2.7.7.48)].
 OS Foot-and-mouth disease virus (strain O1) (Aphthovirus O) (FMDV).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Aphthovirus.
 OX NCBI_TaxID=73482;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OLK;
 RX MEDLINE=84297249; PubMed=6089122;
 RA Fors S., Strebel K., Beck E., Schaller H.;
 RT "Nucleotide sequence and genome organization of foot-and-mouth
 RT disease virus";
 RL Nucleic Acids Res. 12:6587-6601(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O1BFS;
 RX MEDLINE=83143292; PubMed=6298715;
 RA Makoff A.J., Paynter C.A., Rowlands D.J., Boothroyd J.C.;
 RT "Comparison of the amino acid sequence of the major immunogen from
 RT three serotypes of foot and mouth disease virus";
 RL Nucleic Acids Res. 10:8285-8295(1982).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
 RX MEDLINE=89143740; PubMed=2537470;
 RA Acharya R., Fry E., Stuart D., Fox G., Rowlands D., Brown F.;
 RT "The three-dimensional structure of foot-and-mouth disease virus at
 RT 2.9-A resolution.";
 RL Nature 337:709-716(1989).
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate + N diphosphate +
 CC {RNA} (N).
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- MISCELLANEOUS: THE STRAIN O1K SEQUENCE IS SHOWN.
 CC -!- MISCELLANEOUS: THE COAT PROTEIN VP1 CONTAINS THE MAIN ANTIGENIC
 CC DETERMINANTS OF THE VIRION; THEREFORE, CHANGES IN ITS SEQUENCE
 CC MUST BE RESPONSIBLE FOR THE HIGH ANTIGENIC VARIABILITY OF THE
 CC VIRUS.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X00871; CAA25416.1; -;
 CC EMBL; J02185; AAA42635.1; -;
 CC PDB; 1OMY; 18-SEP-01.
 CC PDB; 1OOL; 10-NOV-00.
 CC MEROPS; C03.008; -;
 CC InterPro; IPR004004; Calici_pol_hel.
 CC InterPro; IPR009003; Cys_Ser_trypsin.

DR InterPro; IPR004080; FMDVP1coat.
 DR InterPro; IPR008739; Peptidase_C28.
 DR InterPro; IPR001676; Rhv.
 DR InterPro; IPR000605; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_P5vir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF05408; Peptidase_C28; 1.
 DR Pfam; PF00073; rhv; 3.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICVIRUSNS.
 DR PRINTS; PR01543; FMDVPEICAT.
 KW Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;
 KW Transferase; Hydrolase; Thiol protease; Nonstructural protein;
 KW Myristate; 3D-structure; Lipoprotein.
 FT CHAIN 1 201 NONSTRUCTURAL PROTEIN P20A.
 FT CHAIN 202 286 COAT PROTEIN VP4.
 FT CHAIN 287 504 COAT PROTEIN VP2.
 FT CHAIN 505 724 COAT PROTEIN VP3.
 FT CHAIN 725 937 COAT PROTEIN VP1.
 FT CHAIN 938 1107 CORE PROTEIN P12.
 FT CHAIN 1108 1425 CORE PROTEIN P13.
 FT CHAIN 1426 1578 CORE PROTEIN P14.
 FT CHAIN 1579 1649 GENOME-LINKED PROTEIN VP3.
 FT CHAIN 1650 1862 PROTEASE.
 FT CHAIN 1863 2332 RNA-DIRECTED RNA POLYMERASE.
 FT LIPID 202 202 N-myristoyl glycine (in host).
 FT DISULFID 511 511 INTERCHAIN (IN VP3 DIMER).
 FT DISULFID 406 858 IN VP2-VP1 DIMER.
 FT VARIANT 780 780 I -> V (IN STRAIN O1BFS).
 FT VARIANT 808 808 G -> R (IN STRAIN O1BFS).
 FT VARIANT 861 861 N -> S (IN STRAIN O1BFS).
 SQ SEQUENCE 2332 AA; 258924 MW; 4A83176F43447D68 CRC64;

Query Match 4.5%; Score 129.5; DB 1; Length 2332;

Best Local Similarity 21.2%; Pred. No. 0.24;

Matches 75; Conservative 35; Mismatches 131; Indels 113; Gaps 15;

QY 79 LEIGPELNYLSHLSRMVYAGGMOVVLVAGNAFTAGKIIPAAVPPHPPVENISAQI 138
 Db 367 LELPTDHKVGSLTDSYAMRNGWDVEVTANGNQNGCLLVAMPPELYSIQKRELYQL 426
 QY 139 TMCPHVIVDVRQLEPVLPLPDIR-NRFFHYQENTPRMRLVAMLYTPLRANSGEDVFTV 197
 Db 427 TLEPHQFINPRNTMTAKITVPFVGNRYDOYKHKP--WLVVMVVAPLTVN----- 476
 QY 198 SCRVLTRPADPEFTFLVPPTVESKTPFTLTLTGLSLSNR--PPAID-----MLYT 250
 Db 477 -----TEGAQIKVANIAPT-----NHHVAGEPPSKGIFPVACSDGYGLVTT 521
 QY 251 DPNES-----IVVQPN---GRCITLDTGTTQVLTQVTCACFRGTLISQARAADSTD 301
 Db 522 DKPTADPVYGVKFNPPRNQLPGRFT---NLLDVAEACPT-FLRFEQGVVYVTK---TD 573
 QY 302 SPQARNHPLHVQVKNLDG-----TQYD-----PTD----- 327
 Db 574 SRVLAQFDMSLAAKQMSNTFLAGLAQYVTVQSGTINLHFMFTDAKARYMVAVPPG 633
 QY 328 -----DIPAVLGAIDFKGTGVGVASQRDVSG 353
 Db 634 MBPPTPEAAAHCHIAEMDTGLNSKFTFSIP-YLSAADYVYTAGSVAETTNVQG 686

RESULT 11

POLG_HE701
 ID POLG HE701 STANDARD; PRT; 2194 AA.
 AC P32537;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2

DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein
DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-
DE linked protein VPG (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
OS Human enterovirus 70 (strain J670/71) (EV 70).
OC Viruses; serNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=31915;
RN [1]
RP SEQUENCE FROM N.A. PubMed=2172447;
RX MEDLINE=91037960; PubMed=2172447;
RA Ryan M.D., Jenkins O., Hughes P.J., Brown A., Knowles N.J., Booth D.,
RA Minor P.D., Almond J.W.;
RT "The complete nucleotide sequence of enterovirus type 70:
RT relationships with other members of the picornaviridae.";
RL J. Gen. Virol. 71:2291-2299(1990).
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA) (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: D00820; BAAL891.1; -;
CC FIR: A3253; GNNYE7.
CC HSP: P03300; IPOV.
CC MEROPS: C03_UPE; -;
CC MEROPS: C03_UPE; -;
CC InterPro: IPR003593; AAA_ATPase.
CC InterPro: IPR004004; Calici_pol_hel.
CC InterPro: IPR009003; Cys_Ser_trypsin.
CC InterPro: IPR000199; Pept_3C_picorn.
CC InterPro: IPR000081; Peptidase_C3.
CC InterPro: IPR003138; Pico_P1A.
CC InterPro: IPR002527; Pico_P2B.
CC InterPro: IPR001676; Rhv.
CC InterPro: IPR000605; RNA_helicase.
CC InterPro: IPR007095; RNA_pol_DS_PS.
CC InterPro: IPR001205; RNA_pol_P3D.
CC InterPro: IPR007094; RNA_pol_PSVir.
CC InterPro: IPR008975; Viral_cap_coat.
CC Pfam: PF00548; Cys-protease-3C; 1.
CC Pfam: PF02226; Pico_P1A; 1.
CC Pfam: PF00947; Pico_P2A; 1.
CC Pfam: PF01552; Pico_P2B; 1.
CC Pfam: PF00073; rhv_3.
CC Pfam: PF00680; RNA_dep_RNA_pol; 1.
CC Pfam: PF00910; RNA_helicase; 1.
CC PRINTS: PR00918; CALICVIRUSNS.
CC ProDom: PD001125; Cys_protease_3C; 1.
CC ProDom: PD001306; Pico_P2A; 1.
CC ProDom: PD001274; Pico_P2B; 1.
CC SMART: SM00382; AAA; 1.
KW Polyprotein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
KW Lipoprotein. 2 699 COAT PROTEIN VP4.
FT CHAIN 70 319 COAT PROTEIN VP2.
FT CHAIN 320 561 COAT PROTEIN VP3.

FT CHAIN 871 COAT PROTEIN VP1.
FT CHAIN 872 COAT PROTEIN P2A.
FT CHAIN 1014 CORE PROTEIN P2B.
FT CHAIN 1015 CORE PROTEIN P2C.
FT CHAIN 1113 CORE PROTEIN P2C.
FT CHAIN 1114 CORE PROTEIN P3A.
FT CHAIN 1443 CORE PROTEIN P3A.
FT CHAIN 1444 CORE PROTEIN P3A.
FT CHAIN 1532 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1533 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1534 PICORNAIN 3C.
FT CHAIN 1535 PICORNAIN 3C.
FT CHAIN 1536 RNA-DIRECTED RNA POLYMERASE P3D.
FT CHAIN 1537 RNA-DIRECTED RNA POLYMERASE P3D.
FT CHAIN 1738 N-myristoyl glycine (in host) (BY
FT CHAIN 1739 similarity)
FT ACT SITE 1701 PROTEASE (POTENTIAL).
FT ACT_SITE 1715 PROTEASE (POTENTIAL).
FT ACT_SEQUENCE 2194 AA; 244590 MW; 15DBAE96E06673C CRC64;
SQ SEQUENCE 2194 AA; 244590 MW; 15DBAE96E06673C CRC64;
Query Match 4.3%; Score 125.5; DB 1; Length 2194;
Best Local Similarity 23.3%; Pred. No. 0.44;
Matches 55; Conservative 35; Mismatches 85; Indels 61; Gaps 10;
QY 66 VSPRNSPGEMLLNLEL-----GPELNPYLSHLSRMVNGYAGQVQVVLGNAFTAGKII 120
DB 390 ISAGSDMDQLLEFIPLDILQLEGPLRNTLLGNISRYTHWSGSLDETFMFCGSMFTGKLI 449
QY 121 FAAVPP--HPFVENISAAQITMCPHVIVDVRLPEVLLPLPDIRNRRFFHYNOENTPRML 178
DB 450 ICYTPPGSGSPTDRMQAMLAT---HVWDFGLSSITIIIPWISGS--HYRMENTDAAKAI 504
QY 179 VA-----MLYTPLRANS--EDVFTVSCR-----LTPADDFEFTFLVPP--- 217
DB 505 NANVGYVTCFQTNLAPVGAADQYIVGVAAKDFNLFMRDSDPDIGOSAILPEQAAAT 564
QY 218 -----TVESKTKPF-----TLPTILTGLSELNSRFFPAAIDMLYTDPNESI 256
DB 565 TQIGSEIVKTVANTVESEIKAEGLVIPSINAVEIGATSN-----TEPBBAI 609
RESULT 12
POLG_FMDV1 STANDARD; PRT; 2333 AA.
ID POLG_FMDV1 STANDARD; PRT; 2333 AA.
AC P03306; Q84750; Q84751; Q84752; Q84753; Q84754; Q84760;
AC Q84761; Q84762; Q84763; Q84764; Q84765; Q84766; Q84767; Q84768;
AC Q84769; Q89824;
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last annotation update)
DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat
DE proteins VP1 TO VP4; Core protein P52; Genome-linked proteins VPG1 TO
DE VPG3; Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed
DE RNA polymerase P56A (EC 2.7.7.48)].
DE Foot-and-mouth disease virus (strain A10-61) (Aphthovirus A) (FMDV).
OS Viruses; serNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Aphthovirus.
OC NCBI_TaxID=12112;
RX [1]
RX SEQUENCE FROM N.A. PubMed=6324120;
RX MEDLINE=94169547; PubMed=6324120;
RA Carroll A.R., Rowlands D.J., Clarke B.E.;
RA "The complete nucleotide sequence of the RNA coding for the primary
RA translation product of foot and mouth disease virus.";
RL Nucleic Acids Res. 12:2461-2472(1984).
RN [2]
RN SEQUENCE OF 115-1048 FROM N.A.
RP MEDLINE=82211814; PubMed=6282711;
RX Boothroyd J.C., Harris T.J.R., Rowlands D.J., Lowe P.A.;
RX "The nucleotide sequence of cDNA coding for the structural proteins
RX of foot-and-mouth disease virus.";
RL Gene 17:153-161(1992).
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA) (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.

CC MEDLINE=20556633; PubMed=1102480;
CC Petrecca K., Miller D.M., Shrier A.;
CC "Localization and enhanced current density of the Kv4.2 potassium
CC channel by interaction with the actin-binding protein filamin.";
CC J. Neurosci. 20:8736-8744(2000).
CC [8]
CC INTERACTION WITH INPPL1.
CC MEDLINE=21601813; PubMed=11739414;
CC Dyson J.M., O'Malley C.J., Becanovic J., Munday A.D., Berndt M.C.,
CC Coghill I.D., Nandurkar H.H., Coombs L.M., Mitchell C.A.;
CC "The SH2-containing inositol polyphosphate 5-phosphatase, SHIP-2,
CC binds filamin and regulates submembraneous actin.";
CC J. Cell Biol. 155:1065-1079(2001).
CC [9]
CC REVIEW.
CC MEDLINE=21234905; PubMed=11336782;
CC van der Flier A., Sonnenberg A.;
CC "Structural and functional aspects of filamins.";
CC Biochim. Biophys. Acta 1538:99-117(2001).
CC [10]
CC REVIEW.
CC MEDLINE=21146932; PubMed=11252955;
CC Stoscel T.P., Condeelis J., Cooley L., Hartwig J.H., Noegel A.,
CC Schleicher M., Shapiro S.S.;
CC "Filamins as integrators of cell mechanics and signalling.";
CC Nat. Rev. Mol. Cell Biol. 2:138-145(2001).
CC [11]
CC VARIANTS PH PHE-656 AND THR-1764.
CC MEDLINE=21423782; PubMed=11532987;
CC Sheen V.L., Dixon P.H., Fox J.W., Hong S.E., Kinton L., Sisodiya S.M.,
CC Duncan J.S., Dubeau F., Schaffer I.E., Schachter S.C., Wilner A.,
CC Henchy R., Crino P., Kamuro K., DiMario F., Berg M., Kuzniack R.,
CC Cole A.J., Bromfield E., Biber M., Schomer D., Wheelless J., Silver K.,
CC Mochida G.H., Berkovic S.F., Andermann F., Andermann E., Dobyns W.B.,
CC Wood N.W., Walsh C.A.;
CC "Mutations in the X-linked filamin 1 gene cause periventricular
CC nodular heterotopia in males as well as in females.";
CC Hum. Mol. Genet. 10:1775-1783(2001).
CC [12]
CC VARIANT PH VAL-82.
CC MEDLINE=21912461; PubMed=11914408;
CC Moro F., Carozzo R., Veggioni P., Tortorella G., Toniolo D.,
CC Volzone A., Guerrini R.;
CC "Familial periventricular heterotopia: missense and distal truncating
CC mutations of the FLN1 gene.";
CC Neurology 58:916-921(2002).
CC [13]
CC VARIANTS OPD1 PHE-172; TRP-196 AND LEU-207, VARIANTS OPD2 PRO-170;
CC GLY-196; SER-200; LYS-254; PRO-273; LYS-555 AND PHE-1645, VARIANTS FMD
CC ALA-1159; LEU-1186 AND ILE-1620 DEL, AND VARIANTS MNS GLU-1184;
CC THR-1188 AND LEU-1199.
CC MEDLINE=2252708; PubMed=12612583;
CC Robertson S.P., Twigg S.R.F., Sutherland-Smith A.J., Biancalana V.,
CC Gorlin R.J., Horn D., Kendrick S.J., Kim C.A., Morava E.,
CC Newbury-Ecob R., Oerstavik K.H., Quarrell O.W.J., Schwartz C.E.,
CC Shears D.J., Suri M., Kendrick-Jones J., Wilkie A.O.M.;
CC "Localized mutations in the gene encoding the cytoskeletal protein
CC filamin A cause diverse malformations in humans.";
CC Nat. Genet. 33:487-491(2003).
CC [1]
CC FUNCTION: Promotes orthogonal branching of actin filaments and
CC links actin filaments to membrane glycoproteins. Anchors various
CC transmembrane proteins to the actin cytoskeleton and serves as a
CC scaffold for a wide range of cytoplasmic signaling proteins.
CC Interaction with FLNA may allow neuroblast migration from the
CC intracellular zone into the cortical plate. Tethers cell surface-
CC localized furin, modulates its rate of internalization and directs
CC its intracellular trafficking (By similarity).
CC [2] SUBUNIT: Homodimer. Interacts with FLNB, FURIN, HSPB7, INPPL1,
CC PSEN1 and PSEN2. Interacts also with various other binding
CC partners in addition to filamentous actin. Interacts
CC with KCND2.
CC [3] TISSUE SPECIFICITY: Ubiquitous.
CC [4] SUBCELLULAR LOCATION: Cytoplasmic; peripheral.
CC [5] TISSUE SPECIFICITY: Ubiquitous.

Db 1756 QLAPOQTYAGGQQTWA-----PERPLVGNG-----LDVTS 1787
QY 151 LEPVLLPLDIRRPFYHNOENTPRML-----VAMLYTPLRANSGE- 192
Db 1788 LRPFDLVIP-----FTIKKEITGEVBMPSGKVAQPTITDNKDGTVTVYARSEAGLHEM 1842
QY 193 -----DVTVCRLVTRPAPDFEFTFLVPPTVESKTKPFTLPLTLGEL 236
Db 1843 DIRYDNMHIPGSPLOQYVDYVNGCHVTAYCPG-----LTHGVVKNPATFTVNTKDAGSG 1896
QY 237 SNS---RFPAAIDMLYTDNE-----SIVVQ-----PQNGRCTLG 269
Db 1897 GLSLAIEGSKAEISCTNQDGTCSVSLPVLPGDYSLVLYKNEOHVPGSPPTFARVTGDD 1956
QY 270 TLQ-----GTTQLVPTQICAFRGTLISQARAADSTDSF---QRARN----- 308
Db 1957 SMRSHLKVGSAAIDIPINSETDLSLTATVPPSGREPCLLKRLRNHGVHGISFVPKET 2016
QY 309 --HPLHVQVKNLDGTQYDPTDIPAVLGAIDFKGTGVGVASQDVSQO-OEQGHYATRAH 365
Db 2017 GEHLVHVK-KN---GQHVASSPIPVVISQSE-----IGDASRVVSGQGLHEGHTFEPA- 2066
QY 366 EAHIDTDPKYAPKLTILIKGSD--DENTQPIRFTVPVGMDDNNRWQWELPDYSGRLT 423
Db 2067 EFTIDTRDAGY---GLSLIEGSPSKVDINTED-----LEDGTCRVTYCTFERGNYI 2115
QY 424 LNMNLAPAVSPSPF-----GE---RILFFRSIVPSAGGYSY-IDLCLIPQEWQHF- 471
Db 2116 INIKFADQHVGPSFSPKVTGEORVKESITRRRAPSVANVGHCDLSLKIPEISIQDWT 2175
QY 472 YQEAAPS-----QSAVALRVYVNPDTGRNIFEAKLHREGFLTVCANGNNPIVVP 520
Db 2176 AQVTSPSGKTHEAEIVGEGNHTYCIKRVPAEMGTHTVSVKYGQHV----- 2221
QY 521 PNGYFRFEAGNQFTYLPAMGSG 543
Db 2222 PGSPQF-----TVGPLGEG 2236

RESULT 14

POLG_CX16T STANDARD; PRT; 2193 AA.
ID AC POLG_CX16T
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].
OS Coxsackievirus A16 (strain Tainan/5079/98).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.
OX NCBI_TaxID=231417;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21427116; PubMed=11536241;
RA Yan J.-J., Su I.-J., Chen P.-F., Liu C.-C., Yu C.-K., Wang J.-R.;
RT "Complete genome analysis of enterovirus 71 isolated from an outbreak in Taiwan and rapid identification of enterovirus 71 and coxsackievirus A16 by RT-PCR.";
RL J. Med. Virol. 65:331-339 (2001).
CC -!- FUNCTION: It is thought that the P2C protein attaches to vesicular membranes and is associated with viral RNA synthesis.
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the poliovirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + [RNA] (N).

CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; AF177911; AAC55085.1; -.
DR HSSP; P03300; IPOV.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR004004; Calci_pol_hel.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001999; Pept_3C_picorn.
DR InterPro; IPR000081; Peptidase_C3.
DR InterPro; IPR003138; Pico_P1A.
DR InterPro; IPR002527; Pico_P2B.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00548; Cys-protease_3C; 1.
DR Pfam; PF02226; Pico_P1A; 1.
DR Pfam; PF00947; Pico_P2A; 1.
DR Pfam; PF01552; Pico_P2B; 1.
DR Pfam; PF00073; rhv_3.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICVIRUSNS.
DR ProDom; PD001125; Cys_protease_3C; 1.
DR ProDom; PD001306; Pico_P2A; 1.
DR ProDom; PD001274; Pico_P2B; 1.
DR SMART; SM00382; AAA; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease; Myristate;
KW Lipoprotein.
FT CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 323 COAT PROTEIN VP2.
FT CHAIN 324 565 COAT PROTEIN VP3.
FT CHAIN 566 862 COAT PROTEIN VP1.
FT CHAIN 863 1012 CORE PROTEIN P2A.
FT CHAIN 1013 1111 CORE PROTEIN P2B.
FT CHAIN 1112 1440 CORE PROTEIN P2C.
FT CHAIN 1441 1526 CORE PROTEIN P3A.
FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VP3.
FT CHAIN 1549 1731 PICORNAIN 3C.
FT CHAIN 1732 2193 N-myristoyl glycine (in host) (By similarity).
FT ACT_SITE 1695 1695 PROTEASE (POTENTIAL).
FT ACT_SITE 1709 1709 PROTEASE (POTENTIAL).
SQ SEQUENCE 2193 AA; 243182 MW; 927839DB58F61E7F CRC64;
Query Match 4.1%; Score 118.5; DB 1; Length 2193;
Best Local Similarity 32.5%; Pred. No. 1.5;
Matches 37; Conservative 15; Mismatches 49; Indels 13; Gaps 4;
QY 57 VQAPGGEFTVSPRNSFGEMLLNLEIGPELNPYLSHLRMYNGYAGMQVQVVLGNAFTA 116
Db 398 VQSKTGELCAAPRADPGRD-----GPMQSTILGQLCRYTQWGSGLVTFMFAGSFMAT 451
QY 117 GXIIIAAVPP--HPVENISAQITMCPHVIYDVQLEPVLPLPDIRRFFHY 168
Db 452 GKMLIATPPPGNVNADPITAMLGT---HVIDWDFGLQSSVTLVWPWISNT--HY 500

FT CHAIN 1427 1579 CORE PROTEIN P19.
 FT CHAIN 1580 1602 GENOME-LINKED PROTEIN VP1.
 FT CHAIN 1603 1626 GENOME-LINKED PROTEIN VP2.
 FT CHAIN 1627 1650 GENOME-LINKED PROTEIN VP3.
 FT CHAIN 1651 1863 PROTEASE.
 FT CHAIN 1864 2336 RNA-DIRECTED RNA POLYMERASE.
 FT CHAIN 202 202 N-methylglycine (in host) (By similarity).
 SQ SEQUENCE 2336 AA; 259982 MW; 15AC2AB022B5B954 CRC64;
 Query Match 4.1%; Score 118; DB 1; Length 2336;
 Best Local Similarity 19.0%; Pred. No. 1.8; Indels 82; Gaps 8;
 Matches 67; Conservative 46; Mismatches 158; Indels 82; Gaps 8;
 Qy 64 FTVSPRNSPCGEMLLNLELGPENLPYLSHLSRMVNGVAGGMQVQVVLGNAFTAGKIIFAA 123
 Db 353 FDMTPDKAFGH-LEKLELPTDHHGVYGHVDSFAYMENGWDEVSAGVQNGGCLLVAM 411
 Qy 124 VPHFPVENISAAQITMCPHVIIVRQLBEPVLLPLDIRRPFHYNQENTPMRLVAMLY 183
 Db 412 VPEWKELTPREKYQLTLFPHQFI-----SPRTNMTAHIV 445
 Qy 184 TPLRANGSEDFVTVSCRVLTREPAPDEFELVPTVESKTKPTLPLTLGELNSRFFPA 243
 Db 446 VPYLGVNRYDQY-----KKHKPTLVVMVVSPLTTTVSA 480
 Qy 244 AIDMLYTDPNESIVVQPNQGRCTLDGTLOGTTLVPTQICAFRGTLISQTAARADSDSP 303
 Db 481 GQIKVYAN-----IAPH--VHAGELPSKEGIVPVACSDGYGLVTTDPKTAD----- 527
 Qy 304 QRARNHPLHVQVKNLDGTQY-----DPTDDIPAVLGAIDFKGTGVGVASQRDVSCQ 355
 Db 528 -----PVYGMVYNPPRTNYEGRTNLLDVAEACPTLCFDDGKPYVVTDEQRLAKF 581
 Qy 356 EQGHVATRAHEAHIDTTPKYPAPKLGIT---LIKSGSDDFNNTNPIRFTPVGM 405
 Db 582 DLJSLAAKMSNTYLSGIAQYQAQYSGTINLHFMTGSTDSKARYMVAIVPPGV 634

Search completed: June 1, 2004, 13:47:25
 Job time : 8.64429 secs

RESULT 15
 POLG_FMDVZ STANDARD; PRT; 2336 AA.
 AC P49303; 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat
 DE proteins VP1 TO VP4; Core proteins X, P14, P41, P19; Genome-linked
 DE proteins VP1 TO VP3; Picornain 3C (EC 3.4.22.28) (Protease 3C)
 DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48)].
 OS Foot-and-mouth disease virus (strain A22/550 Azerbaijan 65)
 OS (Aphthovirus A) (FMDV).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Aphthovirus.
 OC NCBI_TaxID=73481;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP Senovtsev S.V., Onischenko A.M., Petrov N.A., Kalaehnikova T.I.,
 RA Mamaeva N.V., Drygin V.Y., Perevozchikova N.A., Vasilenko S.K.;
 RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the
 CC poliovirus polyprotein. In other picornavirus reactions Glu may be
 CC substituted for Gln, and Ser or Thr for Gly.
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA} (N).
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -!- PM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X74812; CAA52812.1; -.
 DR PIR; S37077; S37077.
 DR HSP; O88571; 1TME.
 DR MEROPS; C03.008; -.
 DR MEROPS; C28.001; -.
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR009003; Cys_ser_trypsin.
 DR InterPro; IPR004080; FMDVpicoat.
 DR InterPro; IPR008739; Peptidase_C28.
 DR InterPro; IPR001676; Rhv.
 DR InterPro; IPR006005; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_p3D.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF05408; Peptidase_C28; 1.
 DR Pfam; PF00073; rhv; 3.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICVIRUSNS.
 DR PRINTS; PR01542; FMDVFLCOAT.
 DR PolyProtein; Coat protein; Core protein; RNA-directed RNA polymerase;
 DR Transferase; Hydrolase; Thiol protease; Nonstructural protein;
 DR Myristate; Lipoprotein.
 KW CHAIN 1 201 NONSTRUCTURAL PROTEIN P20A.
 FT CHAIN 202 286 COAT PROTEIN VP4.
 FT CHAIN 287 504 COAT PROTEIN VP2.
 FT CHAIN 505 724 COAT PROTEIN VP3.
 FT CHAIN 725 938 COAT PROTEIN VP1.
 FT CHAIN 939 954 CORE PROTEIN X.
 FT CHAIN 955 1108 CORE PROTEIN P14.
 FT CHAIN 1109 1426 CORE PROTEIN P41.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:35:27 ; Search time 32.0218 Seconds
(without alignments)
5419.275 Million cell updates/sec

Title: US-09-926-799-10
Perfect score: 2906
Sequence: 1 MKMASNDAAPSNDGAANLVP.....GNQFYTLAPMGSGQGRRAQ 550

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2906	100.0	550	12 Q8JW42	Q8JW42 norwalk-lik
2	2861	98.5	550	12 Q318B6	Q318B6 norwalk-lik
3	2855	98.2	550	12 Q318B3	Q318B3 norwalk-lik
4	2847	98.0	550	12 Q91V37	Q91V37 human calic
5	2828	97.3	550	12 Q8V778	Q8V778 norwalk vir
6	2819	97.0	550	12 Q8V781	Q8V781 norwalk vir
7	2637.5	90.8	547	12 Q318A7	Q318A7 norwalk-lik
8	2630.5	90.5	547	12 Q8V775	Q8V775 norwalk vir
9	2623.5	90.3	547	12 Q8V773	Q8V773 norwalk vir
10	2129	73.3	548	12 Q66296	Q66296 calicivirus
11	2129	73.3	548	12 Q8V768	Q8V768 norwalk vir
12	2126	73.2	548	12 Q8V0P5	Q8V0P5 human calic
13	2121	73.0	548	12 Q8V0P4	Q8V0P4 human calic
14	2117	72.8	548	12 Q31722	Q31722 norwalk-lik
15	2117	72.8	548	12 Q91725	Q91725 norwalk-lik
16	2116	72.8	548	12 Q318A1	Q318A1 norwalk-lik

17	2116	72.8	548	12 Q8BC87	Q8BC87 human calic
18	2115	72.8	548	12 Q88291	Q88291 small round
19	2115	72.8	548	12 Q916E5	Q916E5 human calic
20	2114	72.7	548	12 Q68291	Q68291 human calic
21	2114	72.7	548	12 Q80FK9	Q80FK9 human calic
22	2113	72.7	548	12 Q8V771	Q8V771 norwalk vir
23	2113	72.7	548	12 Q8BC90	Q8BC90 human calic
24	2111	72.6	548	12 Q91728	Q91728 norwalk-lik
25	2110	72.6	548	12 Q91V40	Q91V40 human calic
26	2110	72.6	548	12 Q80FK8	Q80FK8 human calic
27	2106	72.5	548	12 Q80FK5	Q80FK5 human calic
28	2105	72.4	548	12 Q83880	Q83880 norwalk vir
29	2104	72.4	548	12 Q80FL1	Q80FL1 human calic
30	2101	72.3	548	12 Q9PYA7	Q9PYA7 human calic
31	2098	72.2	548	12 Q918A4	Q918A4 norwalk-lik
32	2098	72.2	548	12 Q80FK7	Q80FK7 human calic
33	2097	72.2	548	12 Q80FK6	Q80FK6 human calic
34	2096	72.1	548	12 Q80FK1	Q80FK1 human calic
35	2095	72.1	548	12 Q80FK3	Q80FK3 human calic
36	2093	72.0	548	12 Q80FK4	Q80FK4 human calic
37	2090	71.9	548	12 Q915D3	Q915D3 human calic
38	2088	71.9	548	12 Q915D2	Q915D2 human calic
39	2083	71.7	548	12 Q31V46	Q31V46 human calic
40	2078	71.5	548	12 Q80FK2	Q80FK2 human calic
41	2077	71.5	548	12 Q80FL2	Q80FL2 human calic
42	2074	71.4	548	12 Q80FL0	Q80FL0 human calic
43	2064.5	71.0	535	12 Q911I5	Q911I5 human calic
44	2062.5	71.0	535	12 Q9QMK6	Q9QMK6 chitta viru
45	2059.5	70.9	535	12 Q915C6	Q915C6 human calic

ALIGNMENTS

RESULT 1

Q8JW42 Q8JW42 PRELIMINARY; PRT; 550 AA.
AC Q8JW42;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
GN CAPSID.
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ueno7k;
RA Katayama K., Takeda N., Natori K.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ueno7k;
RA Natori K., Takeda N.;
RT "Genetic and antigenic relationship among Norwalk-like viruses.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB078337; BAC05518.1;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 550 AA; 60321 MW; 543265E29E05D1F0 CRC64;

Query Match 100.0%; Score 2906; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 1.3e-227;
Matches 550; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKMASNDAAPSNDGAANLVPENDEVMALPEVVGASIAAPVVGQNIIDPWIRENFVQAP 60
DB 1 MKMASNDAAPSNDGAANLVPENDEVMALPEVVGASIAAPVVGQNIIDPWIRENFVQAP 60
QY 61 QGEFTVSPRNSPGEMLLNLGLPELNPYLHLSRMVNGYAGQMVGQVVLGNAGNFTAGKII 120

DR Pfam; PF00915; Calici_coat; 1. 4B8AE1352B79D995 CRC64;
SQ SEQUENCE 550 AA; 60374 MW; 98.5% Score 2861; DB 12; Length 550;
Query Match 98.7%; Pred. No. 6.1e-224; Indels 0; Gaps 0;
Best Local Similarity 98.7%; Mismatches 1; Conservative 543; Conserved 543;
Matches 543; Conserved 543;
Qy 1 MKVASNDAPSDGAAANLVPEANDEWMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAP 60
Db 1 MKVASNDAPSDGAAANLVPEANDEWMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAP 60
Qy 61 QGEFTVSPRNSPGEMLLNLELGPENLPYLSHLSRMVYAGGMOVVLGAGNAFTAGKII 120
Db 61 QGEFTVSPRNSPGEMLLNLELGPENLPYLSHLSRMVYAGGMOVVLGAGNAFTAGKII 120
Qy 121 FPAADMLYTDPNESIVVQPNQGRCTLDGTLQGTTLQVPTQICAFRGTLISQTAAADST 300
Db 121 FPAADMLYTDPNESIVVQPNQGRCTLDGTLQGTTLQVPTQICAFRGTLISQTAAADST 300
Qy 181 MLYTPLRANGSGEDVFTVSCRVLTRPADPDEFTFLVPPPTVESKTKPFTLPTLTGELSNR 240
Db 181 MLYTPLRANGSGEDVFTVSCRVLTRPADPDEFTFLVPPPTVESKTKPFTLPTLTGELSNR 240
Qy 241 FPAADMLYTDPNESIVVQPNQGRCTLDGTLQGTTLQVPTQICAFRGTLISQTAAADST 300
Db 241 FPAADMLYTDPNESIVVQPNQGRCTLDGTLQGTTLQVPTQICAFRGTLISQTAAADST 300
Qy 301 DSPQARARNHPLHVQKNLDGTYDPTDDIPAVLGAIDFKGTGVSQSDVSGQEQGHY 360
Db 301 DSPQARARNHPLHVQKNLDGTYDPTDDIPAVLGAIDFKGTGVSQSDVSGQEQGHY 360
Qy 361 ATRAHAHIDTDPKAPKLTGTLIKSGSDDFNTNQPPIRFTPVGMGNNRQWELPDYSG 420
Db 361 ATRAHAHIDTDPKAPKLTGTLIKSGSDDFNTNQPPIRFTPVGMGNNRQWELPDYSG 420
Qy 421 RLTLNMLAPAVSPSPGGERILFRSIVPSAGGYSYDCLIPQEWVQHFYQEAAPSQS 480
Db 421 RLTLNMLAPAVSPSPGGERILFRSIVPSAGGYSYDCLIPQEWVQHFYQEAAPSQS 480
Qy 481 AVALRVYVNPDTGRNIFEAHLHREGFLTVANGNNPIVVPNGYFRFPAWGNQFYTLAPM 540
Db 481 AVALRVYVNPDTGRNIFEAHLHREGFLTVANGNNPIVVPNGYFRFPAWGNQFYTLAPM 540
Qy 541 GSGQGRRAQ 550
Db 541 GSGQGRRAQ 550
RESULT 3
Q918B3 ID Q918B3 PRELIMINARY; PRT; 550 AA.
AC Q918B3; SEQUENCE FROM N.A.
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Florida/269/1993/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OC NCBI_TaxID=171839;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Hu/NLV/Florida/269/1993/US;
RC MEDLINE=97193806; PubMed=9041391;
RX Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RA "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Florida/269/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Florida/269/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Florida/269/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Florida/269/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Florida/269/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Florida/269/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Florida/269/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Florida/269/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Florida/269/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Florida/269/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Florida/269/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Florida/269/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Florida/269/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Florida/269/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Florida/269/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Florida/269/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Florida/269/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Florida/269/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., F

RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:8336-8348 (2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/Baltimore/274/1993/US;
RA Ando T., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414408; AAL12977.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 550 AA; 60300 MW; 843401DE6FFBC055 CRC64;

Query Match 98.2%; Score 2855; DB 12; Length 550;
Best Local Similarity 98.5%; Pred. No. 1.9e-223;
Matches 542; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 1 MKMASNDAAAPSDGAANLVPEANDEVMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAP 60
DB 1 MKMASNDAAAPSDGAANLVPEANDEVMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAP 60
QY 61 QGEFTVSPNSPCEMLLNLELGPENLPYLSHLSRMVYAGGVQVVLGNAFTAGKII 120
DB 61 QGEFTVSPNSPCEMLLNLELGPENLPYLSHLSRMVYAGGVQVVLGNAFTAGKII 120
QY 121 FFAVPPHPPVENISAAQITMCCHVIVDVRLQLEPVLLPDIRNRFHYNQENTPRMRLVA 180
DB 121 FFAVPPHPPVENISAAQITMCCHVIVDVRLQLEPVLLPDIRNRFHYNQENTPRMRLVA 180
QY 181 MLYTPLRANGSDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPTLPTLTGELSNSR 240
DB 181 MLYTPLRANGSDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPTLPTLTGELSNSR 240
QY 241 FFAAIDMLYTDNIESVWQPONGRCTLDGTQTTQTLVPTQICAFRGTLISQTARAADST 300
DB 241 FFAAIDMLYTDNIESVWQPONGRCTLDGTQTTQTLVPTQICAFRGTLISQTARAADST 300
QY 301 DSPQARARNHPLHVQVKNLQDGTQVDDTDIPAVLGAIDFKGTGFGVASQRDVSGQOQGHY 360
DB 301 DSPQARARNHPLHVQVKNLQDGTQVDDTDIPAVLGAIDFKGTGFGVASQRDVSGQOQGHY 360
QY 361 ATRAHAHIDTDPKYPAPKLGTLIKSGDDDFNTNQPIRFTPVGMGDNWNRQWELPDYSG 420
DB 361 ATRAHAHIDTDPKYPAPKLGTLIKSGDDDFNTNQPIRFTPVGMGDNWNRQWELPDYSG 420
QY 421 RLTLNNNLAPAVSPSPGGERILFFRSIVPSAGYSGYIDCLIPQEWVQHFYQEAAPQS 480
DB 421 RLTLNNNLAPAVSPSPGGERILFFRSIVPSAGYSGYIDCLIPQEWVQHFYQEAAPQS 480
QY 481 AVALVRYVNPDTGRNIFEAKLHREGFLTVANGCNPNPIVPPNGYFRFEAWGNQFYTLAPM 540
DB 481 AVALVRYVNPDTGRNIFEAKLHREGFLTVANGCNPNPIVPPNGYFRFEAWGNQFYTLAPM 540
QY 541 GSGQGRRAQ 550
DB 541 GSGQGRRAQ 550

RESULT 4
Q9IV37 PRELIMINARY; PRT; 550 AA.
AC Q9IV37;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Seacroft/90/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=122924;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/Seacroft/90/UK;
RX MEDLINE=20404883; PubMed=10949950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Viruses Genes 20:227-236 (2000).
DR EMBL; AJ277620; CAB89101.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 550 AA; 60355 MW; 71AF0C5CA1DA0506 CRC64;

Query Match 98.0%; Score 2847; DB 12; Length 550;
Best Local Similarity 98.4%; Pred. No. 8.4e-223;
Matches 541; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 1 MKMASNDAAAPSDGAANLVPEANDEVMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAP 60
DB 1 MKMASNDAAAPSDGAANLVPEANDEVMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAP 60
QY 61 QGEFTVSPNSPCEMLLNLELGPENLPYLSHLSRMVYAGGVQVVLGNAFTAGKII 120
DB 61 QGEFTVSPNSPCEMLLNLELGPENLPYLSHLSRMVYAGGVQVVLGNAFTAGKII 120
QY 121 FFAVPPHPPVENISAAQITMCCHVIVDVRLQLEPVLLPDIRNRFHYNQENTPRMRLVA 180
DB 121 FFAVPPHPPVENISAAQITMCCHVIVDVRLQLEPVLLPDIRNRFHYNQENTPRMRLVA 180
QY 181 MLYTPLRANGSDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPTLPTLTGELSNSR 240
DB 181 MLYTPLRANGSDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPTLPTLTGELSNSR 240
QY 241 FFAAIDMLYTDNIESVWQPONGRCTLDGTQTTQTLVPTQICAFRGTLISQTARAADST 300
DB 241 FFAAIDMLYTDNIESVWQPONGRCTLDGTQTTQTLVPTQICAFRGTLISQTARAADST 300
QY 301 DSPQARARNHPLHVQVKNLQDGTQVDDTDIPAVLGAIDFKGTGFGVASQRDVSGQOQGHY 360
DB 301 DSPQARARNHPLHVQVKNLQDGTQVDDTDIPAVLGAIDFKGTGFGVASQRDVSGQOQGHY 360
QY 361 ATRAHAHIDTDPKYPAPKLGTLIKSGDDDFNTNQPIRFTPVGMGDNWNRQWELPDYSG 420
DB 361 ATRAHAHIDTDPKYPAPKLGTLIKSGDDDFNTNQPIRFTPVGMGDNWNRQWELPDYSG 420
QY 421 RLTLNNNLAPAVSPSPGGERILFFRSIVPSAGYSGYIDCLIPQEWVQHFYQEAAPQS 480
DB 421 RLTLNNNLAPAVSPSPGGERILFFRSIVPSAGYSGYIDCLIPQEWVQHFYQEAAPQS 480
QY 481 AVALVRYVNPDTGRNIFEAKLHREGFLTVANGCNPNPIVPPNGYFRFEAWGNQFYTLAPM 540
DB 481 AVALVRYVNPDTGRNIFEAKLHREGFLTVANGCNPNPIVPPNGYFRFEAWGNQFYTLAPM 540
QY 541 GSGQGRRAQ 550
DB 541 GSGQGRRAQ 550

RESULT 5
Q8V778 PRELIMINARY; PRT; 550 AA.
AC Q8V778;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF2 protein (Capsid).
OS Norwalk-like virus.
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=11983, 95340;
RN [1]
RP SEQUENCE FROM N.A.


```

QY 421 RLTLNMLNLA VSPSPGGERILFFRSIVPSAGGSGYIDCLIPQEWVQHFYQEAAPSQS 480
DB 421 RLALNMLNLA VSPSPGGERILFFRSIVPSAGGSGYIDCLIPQEWVQHFYQEAAPSQS 480
QY 481 AVALVRYNPDGRNIFFAKLHREGFLTVANGNNPIVVPNGYFRFEAWGNQFTYLLAPM 540
DB 481 AVALVRYNPDGRNIFFAKLHREGFLTVANGNNPIVVPNGYFRFEAWGNQFTYLLAPM 540
QY 541 GSGQGRRAQ 550
DB 541 GSGQGRRAQ 550

RESULT 7
Q918A7
ID Q918A7 PRELIMINARY; PRT; 547 AA.
AC Q918A7;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Miami/292/1994/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171842;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Miami/292/1994/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Miami/292/1994/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
RA Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995.";
RL J. Med. Virol. 53:372-383(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Miami/292/1994/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Miami/292/1994/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414410; AAL12983.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 547 AA; 59931 MW; 5PBEECCA65BA5AEA CRC64;

Query Match 90.8%; Score 2637.5; DB 12; Length 547;
Best Local Similarity 90.0%; Pred. No. 9.1e-206;
Matches 495; Conservative 27; Mismatches 25; Indels 3; Gaps 1;

QY 1 MKMASNDRAAPSNDGAANLVPEANDEVALEPVVGASIAAPVVGQNIIDPWIRENFVQAP 60
DB 1 MKMASNDRAAPSNDGAANLVPEANDEVALEPVVGASIAAPVVGQNIIDPWIRENFVQAP 60
QY 61 QGEFTVSPRNPGEMLLNLELGPENLPSYLSHLSRMNGYAGGMQVQVVLGNAFTAGKII 120

```

```

DB 61 QGEFTVSPRNPGEMLLNLELGPENLPSYLSHLSRMNGYAGGMQVQVVLGNAFTAGKII 120
QY 121 FAAPPPHPVENISAAQITMCPHVIQVDRQLEPVLPLPDINRRFFHYNQENTPMRLVA 180
DB 121 FAAPPPHPVENISAAQITMCPHVIQVDRQLEPVLPLPDINRRFFHYNQENTPMRLVA 180
QY 181 MLYTPLRANSGEDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKDFTLPLTLGELSRSR 240
DB 181 MLYTPLRANSGEDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKDFTLPLTLGELSRSR 240
QY 241 PPAIDMLYTDPNESIVVQPNQRCCTLDGTLOGTTLQVPTQICAFRGTLISQTAADST 300
DB 241 PPAIDMLYTDPNESIVVQPNQRCCTLDGTLOGTTLQVPTQICAFRGTLISQTAADST 300
QY 301 DSPQARNHPLHVQKNLDGTQYDPTDDIPAVLGAIDFKGTGTVFVGSORDVSGOQEGHY 360
DB 301 DSPQARNHPLHVQKNLDGTQYDPTDDIPAVLGAIDFKGTGTVFVGSORDVSGOQEGHY 360
QY 357 301 DSAPARNHPLHVQLKNDLGTQYDPTDEVPVAVLGAIDFKGTGTVFVGSORNTTSSVG--- 357
QY 420 361 ATRAHAHIDTDPKYAPKLGTILIKSGDDFNNTQPIRFTPVGMGDNRMWQELPDYSG 420
DB 420 361 ATRAHAHIDTDPKYAPKLGTILIKSGDDFNNTQPIRFTPVGMGDNRMWQELPDYSG 420
QY 417 358 ATRAHEVHIDTDPKYAPKLGTILIKSGDDFNNTQPIRFTPVGMGDNRMWQELPDYSG 417
DB 417 358 ATRAHEVHIDTDPKYAPKLGTILIKSGDDFNNTQPIRFTPVGMGDNRMWQELPDYSG 417
QY 480 421 RLTLNMLNLA VSPSPGGERILFFRSIVPSAGGSGYIDCLIPQEWVQHFYQEAAPSQS 480
DB 480 421 RLTLNMLNLA VSPSPGGERILFFRSIVPSAGGSGYIDCLIPQEWVQHFYQEAAPSQS 480
QY 540 478 AVALVRYNPDGRNIFFAKLHREGFLTVANGNNPIVVPNGYFRFEAWGNQFTYLLAPM 540
DB 540 478 AVALVRYNPDGRNIFFAKLHREGFLTVANGNNPIVVPNGYFRFEAWGNQFTYLLAPM 540
QY 537 541 GSGQGRRAQ 550
DB 537 541 GSGQGRRAQ 550
DB 547 538 GTGQGRRAQ 547

RESULT 8
Q8V775
ID Q8V775 PRELIMINARY; PRT; 547 AA.
AC Q8V775;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE ORF2 protein (Capsid).
OS Norwalk virus, and
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=11983, 95340;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Norwalk virus; STRAIN=U16GII;
RA Kojima S., Kageyama T., Fukushi S., Hoshino F.B., Shinohara M.,
RA Uchida K., Natori K., Takeda N., Katayama K.;
RT "Genogroup-specific primers for detect Norwalk like virus.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Norwalk-like virus; STRAIN=Saitama U16;
RA Katayama K.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=Norwalk-like virus; STRAIN=Saitama U16;
RX MEDLINE=22192455; PubMed=12202225;
RA Katayama K., Shirato-Horikoshi H., Kojima S., Kageyama T., Oka T.,
RA Hoshino F.B., Fukushi S., Shinohara M., Uchida K., Suzuki Y.,
RA Gojobori T., Takeda N.;
RT "Phylogenetic Analysis of the Complete Genome of 18 Norwalk-like
RT Viruses.";
RL Virology 299:225-239(2002).
DR EMBL; AB067539; BAB84146.1; -.
DR EMBL; AB039778; BAC11825.1; -.

```



```

RA Katayama K.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SPECIES=Norwalk-like virus; STRAIN=Saitama U17;
RC MEDLINE=22192455; PubMed=12202225;
RX Katayama K., Shitoto-Horikoshi H., Kojima S., Kageyama T., Oka T.,
RA Hosono F.B., Fukushi S., Shinohara M., Uchida K., Suzuki Y.,
RA Gojodori T., Takeda N.;
RT "Phylogenetic Analysis of the Complete Genome of 18 Norwalk-like
RT Viruses.";
RL Virology 299:225-239 (2002).
DR EMBL; AB07540; BAC11828.1; --
DR ENBL; AB039779; BAC11828.1; --
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 547 AA; 59916 MW; 5C4F7272208B62B6 CRC64;

Query Match 90.5%; Score 2630.5; DB 12; Length 547;
Best Local Similarity 89.8%; Pred. No. 3.4e-205;
Matches 494; Conservative 27; Mismatches 26; Indels 3; Gaps 1;

QY 1 MKMASNDAAPSNDGAANLVPEANDEVMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAP 60
Db 1 MKMASNDAAPSNDGAANLVPEANDEVMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAP 60
QY 61 QGEFTVSPNSGEMLLNLELGPENLPYLHSLRMYNGVAGQVQVVLGNAFTAGKII 120
Db 61 QGEFTVSPNSGEMLLNLELGPENLPYLHSLRMYNGVAGQVQVVLGNAFTAGKII 120
QY 121 FFAVPPHFPVENISAAQITMCPHVIDVRLQLEPVLPLDIRNRFHYNQETPRMRLVA 180
Db 121 FFAVPPHFPVENISAAQITMCPHVIDVRLQLEPVLPLDIRNRFHYNQETPRMRLVA 180
QY 181 MLYTPLRANGSDVFTVSCRVLTRPAPDFEFTLVPTVESKTKPFTLPLTLGELSNSR 240
Db 181 MLYTPLRANGSDVFTVSCRVLTRPAPDFEFTLVPTVESKTKPFTLPLTLGELSNSR 240
QY 241 FPAIDMLYTDPNESIVVQPNQGRCTLDGTTLQTTQVPTQICSFRTGLISQTSRADST 300
Db 241 FPAIDMLYTDPNESIVVQPNQGRCTLDGTTLQTTQVPTQICSFRTGLISQTSRADST 300
QY 301 DSPQARNHPLHVQKNLDGTQDTPDDIPAVLGAIDFKGTGFGVASQRDVSGQOEGHY 360
Db 301 DSPQARNHPLHVQKNLDGTQDTPDDIPAVLGAIDFKGTGFGVASQRDVSGQOEGHY 360
QY 361 ATRAHEAHIDTDPKYAPKLTILIKSGSDDFNTNPIRFTPVGMGNWQWELPDYSG 420
Db 361 ATRAHEAHIDTDPKYAPKLTILIKSGSDDFNTNPIRFTPVGMGNWQWELPDYSG 420
QY 358 ATRAHEAHIDTTPRYTPKLSVLMYSSDFDDGQPTFTPIGMGADWHQWELPEYSG 417
Db 358 ATRAHEAHIDTTPRYTPKLSVLMYSSDFDDGQPTFTPIGMGADWHQWELPEYSG 417
QY 421 RLTLNMLAPAVSPFGERILFFRSIVPSAGGSGYIDCLIPQEWVQHFYQEAAPSQS 480
Db 421 RLTLNMLAPAVSPFGERILFFRSIVPSAGGSGYIDCLIPQEWVQHFYQEAAPSQS 480
QY 418 HLTLMNMLAPAVAPAFGERILFFRSIVPSAGGSGYIDCLIPQEWVQHFYQEAAPSQS 477
Db 418 HLTLMNMLAPAVAPAFGERILFFRSIVPSAGGSGYIDCLIPQEWVQHFYQEAAPSQS 477
QY 481 AVALRVYNPDGTGRNIFEAKLHREGFLTVANGNNPIVPPNGYFRFEAWNQFYTLAPM 540
Db 481 AVALRVYNPDGTGRNIFEAKLHREGFLTVANGNNPIVPPNGYFRFEAWNQFYTLAPM 540
QY 478 AVALRVYNPDGTGRNIFEAKLHREGFLTVANGNNPIVPPNGYFRFEAWNQFYTLAPM 537
Db 478 AVALRVYNPDGTGRNIFEAKLHREGFLTVANGNNPIVPPNGYFRFEAWNQFYTLAPM 537
QY 541 GSGQGRRAQ 550
Db 541 GSGQGRRAQ 550
QY 538 GTGQGRRAQ 547
Db 538 GTGQGRRAQ 547

RESULT 9
Q8V773 ID Q8V773 PRELIMINARY; PRT; 547 AA.
AC Q8V773
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF2 protein (Capsid).
OS Norwalk virus, and
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=11983, 95340;
RN [1]
RP SPECIES=Norwalk virus; STRAIN=U17GII;
RC SEQUENCE FROM N.A.
RA Kojima S., Kageyama T., Fukushi S., Hoshino F.B., Shinohara M.,
RA Uchida K., Natori K., Takeda N., Katayama K.;
RT "Genogroup-specific primers for detect Norwalk like virus.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP SPECIES=Norwalk-like virus; STRAIN=Saitama U17;
RC SEQUENCE FROM N.A.

```



```
RESULT 12
Q8VOP5 PRELIMINARY; PRT; 548 AA.
ID Q8VOP5;
AC Q8VOP5;
DT 01-MAR-2002 (TREMREL. 20, Created)
DT 01-MAR-2002 (TREMREL. 20, Last sequence update)
DT 01-OCT-2003 (TREMREL. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus Hu/NLV/GII/MD101-2/1987/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OC NCBI_TaxID=159311;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=Hu/NLV/GII/MD101-2/1987/US;
RX MEDLINE=21666333; PubMed=11807686;
RA Green K.Y., Belliot G., Taylor J.L., Valdesuso J., Lew J.F.,
RA Kapikian A.Z., Lin F.Y.C.;
RA "A Predominant Role for Norwalk-like Viruses as Agents of Epidemic
RT Gastroenteritis in Maryland Nursing Homes for the Elderly.";
RL J. Infect. Dis. 185:133-146(2002).
DR EMBL; AY030313; AAK54360.1; -
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR SEQUENCE 548 AA; 53698 MW; 776F74B4E58B6CC7 CRC64;
SQ
Query Match 73.2%; Score 2126; DB 12; Length 548;
Best Local Similarity 71.5%; Pred. No. 3.6e-164;
Matches 397; Conservative 68; Mismatches 78; Indels 12; Gaps 4;
Qy 1 MKMASNDAAPSNDGAANLVPEANDEVMALPEVVGASIAAPVVGQNIIDPWIRENFVQAP 60
Db 1 MKMASNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAFLTQQNIIDPWIMNFVQAP 60
Qy 61 QGEFTVSPRSPGEMLNLELGPPEINLYLHLSRMVYAGGQVQVVLGNAFTAGKII 120
Db 61 QGEFTVSPRSPGEMLNLELGPPEINLYLHLSRMVYAGGQVQVVLGNAFTAGKVI 120
Qy 61 QGEFTVSPRSPGEMLNLELGPPEINLYLHLSRMVYAGGQVQVVLGNAFTAGKII 120
Db 61 QGEFTVSPRSPGEMLNLELGPPEINLYLHLSRMVYAGGQVQVVLGNAFTAGKVI 120
Qy 121 FAAPVPPHPPVENISAQITMCPHVIVDVRQLEPVLPLDPIRNFHYNGENTPRMLVA 180
Db 121 FAAPVPPHPPVENISAQITMCPHVIVDVRQLEPVLPLDPIRNFHYNGENTPRMLVA 180
Qy 121 FAAPVPPHPPVENISAQITMCPHVIVDVRQLEPVLPLDPIRNFHYNGENTPRMLVA 180
Db 121 FAAPVPPHPPVENISAQITMCPHVIVDVRQLEPVLPLDPIRNFHYNGENTPRMLVA 180
Qy 181 MLYTPLRA-NGEDVFTVSCRVLTRPAPDEFTFLVPPTVESKTKPTLPLITISEMSNS 239
Db 181 MLYTPLRA-NGEDVFTVSCRVLTRPAPDEFTFLVPPTVESKTKPTLPLITISEMSNS 240
Qy 181 MLYTPLRANSGDDVFTVSCRVLTRPAPDEFTFLVPPTVESKTKPTLPLITISEMSNS 240
Db 181 MLYTPLRANSGDDVFTVSCRVLTRPAPDEFTFLVPPTVESKTKPTLPLITISEMSNS 240
Qy 240 RPPAIDMLYTDPNESIIVVQNGRCITLDGTLQGTTLQVPTQICAFRGTLISQARAADS 299
Db 241 RPPVIDSLTSPTESIVVQNGRCITLDGTLQGTTLQVPTQICAFRGTLISQARAADS 300
Qy 300 TD-SFQARNHPLHVQVKNLQGTQVDDTDDIPAVLGADFKGTGFGVASQRDVSGQEQ 358
Db 301 ADTATPRLFNHWHIQLDNLNGTTPYDPAEDIPAPLGTPDFRGKVGFGVASQRPDS 355
Qy 359 HYATRAHEAHITDPKVPKLGTLIKSGDDFNTPQIRFTPVGMGDN--WRQWEL 415
Db 356 --TTAHEAKVDTTSGRFTPKLGSLETTESDDFPNQSTKFTPTVIGVGDNEADFQW 413
Qy 416 PDYSGRLTNMNLAPAVSPFPERILFPRSVIPSGAGGVSIGYIDCLIPQEWVQHYQEA 475
Db 414 PDYSGOFTNMNLAPAVAFNPFQEQQLLFFRSQLPSSGGRSNGILDCLVPQEWVQHYQES 473
Qy 476 APSQSAVALRVYNPDTGRNIPEAKLHREGFTLVANCNNPIVPPNGYFRFEANGNOFY 535
Db 474 APAQTQVALRVYNPDTGRVLEAKLHKLGFMTIAKNGDSPTVPPNGYFRFESWVNP 533
Qy 536 TLAPMGSGQGRRAQ 550
Db 534 TLAPMGTGNGRRRIQ 548

RESULT 13
Q8VOP4 PRELIMINARY; PRT; 548 AA.
ID Q8VOP4;
AC Q8VOP4;
DT 01-MAR-2002 (TREMREL. 20, Created)
DT 01-MAR-2002 (TREMREL. 20, Last sequence update)
DT 01-OCT-2003 (TREMREL. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus Hu/NLV/GII/MD134-10/1987/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OC NCBI_TaxID=159312;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=Hu/NLV/GII/MD134-10/1987/US;
RX MEDLINE=21666333; PubMed=11807686;
RA Green K.Y., Belliot G., Taylor J.L., Valdesuso J., Lew J.F.,
RA Kapikian A.Z., Lin F.Y.C.;
RA "A Predominant Role for Norwalk-like Viruses as Agents of Epidemic
RT Gastroenteritis in Maryland Nursing Homes for the Elderly.";
RL J. Infect. Dis. 185:133-146(2002).
DR EMBL; AY030313; AAK54360.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR SEQUENCE 548 AA; 53899 MW; 0C02C5D953500351 CRC64;
SQ
Query Match 73.0%; Score 2121; DB 12; Length 548;
Best Local Similarity 71.4%; Pred. No. 9.2e-164;
Matches 396; Conservative 69; Mismatches 78; Indels 12; Gaps 4;
Qy 1 MKMASNDAAPSNDGAANLVPEANDEVMALPEVVGASIAAPVVGQNIIDPWIRENFVQAP 60
Db 1 MKMASNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAFLTQQNIIDPWIMNFVQAP 60
Qy 61 QGEFTVSPRSPGEMLNLELGPPEINLYLHLSRMVYAGGQVQVVLGNAFTAGKII 120
Db 61 QGEFTVSPRSPGEMLNLELGPPEINLYLHLSRMVYAGGQVQVVLGNAFTAGKVI 120
Qy 61 QGEFTVSPRSPGEMLNLELGPPEINLYLHLSRMVYAGGQVQVVLGNAFTAGKII 120
Db 61 QGEFTVSPRSPGEMLNLELGPPEINLYLHLSRMVYAGGQVQVVLGNAFTAGKVI 120
Qy 121 FAAPVPPHPPVENISAQITMCPHVIVDVRQLEPVLPLDPIRNFHYNGENTPRMLVA 180
Db 121 FAAPVPPHPPVENISAQITMCPHVIVDVRQLEPVLPLDPIRNFHYNGENTPRMLVA 180
Qy 121 FAAPVPPHPPVENISAQITMCPHVIVDVRQLEPVLPLDPIRNFHYNGENTPRMLVA 180
Db 121 FAAPVPPHPPVENISAQITMCPHVIVDVRQLEPVLPLDPIRNFHYNGENTPRMLVA 180
Qy 181 MLYTPLRA-NGEDVFTVSCRVLTRPAPDEFTFLVPPTVESKTKPTLPLITISEMSNS 239
Db 181 MLYTPLRANSGDDVFTVSCRVLTRPAPDEFTFLVPPTVESKTKPTLPLITISEMSNS 240
Qy 240 RPPAIDMLYTDPNESIIVVQNGRCITLDGTLQGTTLQVPTQICAFRGTLISQARAADS 299
Db 241 RPPVIDSLTSPTESIVVQNGRCITLDGTLQGTTLQVPTQICAFRGTLISQARAADS 300
Qy 300 TD-SFQARNHPLHVQVKNLQGTQVDDTDDIPAVLGADFKGTGFGVASQRDVSGQEQ 358
Db 301 ADTATPRLFNHWHIQLDNLNGTTPYDPAEDIPAPLGTPDFRGKVGFGVASQRPDS 355
Qy 359 HYATRAHEAHITDPKVPKLGTLIKSGDDFNTPQIRFTPVGMGDN--WRQWEL 415
Db 356 --TTAHEAKVDTTSGRFTPKLGSLETTESDDFPNQSTKFTPTVIGVGDNEADFQW 413
Qy 416 PDYSGRLTNMNLAPAVSPFPERILFPRSVIPSGAGGVSIGYIDCLIPQEWVQHYQEA 475
Db 414 PDYSGOFTNMNLAPAVAFNPFQEQQLLFFRSQLPSSGGRSNGILDCLVPQEWVQHYQES 473
Qy 476 APSQSAVALRVYNPDTGRNIPEAKLHREGFTLVANCNNPIVPPNGYFRFEANGNOFY 535
Db 474 APAQTQVALRVYNPDTGRVLEAKLHKLGFMTIAKNGDSPTVPPNGYFRFESWVNP 533
Qy 536 TLAPMGSGQGRRAQ 550
Db 534 TLAPMGTGNGRRRIQ 548
```

```

RESULT 14
Q91722 PRELIMINARY; PRT; 548 AA.
ID Q91722;
AC Q91722;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Brattleboro/321/1995/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Brattleboro/321/1995/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Brattleboro/321/1995/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
RA Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995.";
RL J. Med. Virol. 53:372-383(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Brattleboro/321/1995/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Brattleboro/321/1995/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414415; AAL12998.1; --
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 59933 MW; 7B9AFF9AF1469158 CRC64;

Query Match 72.8%; Score 2117; DB 12; Length 548;
Best Local Similarity 71.2%; Pred. No. 1.9e-163;
Matches 395; Conservative 71; Mismatches 77; Indels 12; Gaps 4;

QY 1 MKWASNDAAPSNDGAANLVPEANDEVMALEPVVGSIAAPVVGQNNIIDPWIRENFVQAP 60
Db 1 MKWASNDAAPSNDGAAGLVPEINNEAMALEPVAGAIAAPLTGQQNIIDPWIMNWFVQAP 60
QY 61 QGEFTVSPNSGEMLLNLELPELNPYLHSLRMVYAGQVQVVLGNAFTAGKI 120
Db 61 QGEFTVSPNSGEMLLNLELPELNPYLHSLRMVYAGQVQVVLGNAFTAGKI 120
QY 121 FAAVPHPPVENISAQITMCQPHVIVDVRQLEPVLPLDIERPFFHYNQENTPRMLVA 180
Db 121 FAAIPNFPIDNLAAQITMCQPHVIVDVRQLEPVLPLDIERPFFHYNQENTPRMLVA 180
QY 181 MLYTPLRA-NSGEDVFTVSCRVLTTPADFEFTFLVPVTSKTPFTPLITLGEISNS 239
Db 181 MLYTPLRANSGDDVFTVSCRVLTTPSPDFSNFLVPVTSKTKLFTPLITLISEMSNS 240
QY 240 RPPAIDMLYTPDPSIVVQPNQRCCTLDGTLOGTTLQVPTQICAFRGLTISQTRAAADS 299
Db 240 RPPAIDMLYTPDPSIVVQPNQRCCTLDGTLOGTTLQVPTQICAFRGLTISQTRAAADS 299

```

```

Db 241 RFPVIDSLHTSPTEINIVQCONGRVTLDELMTGTTQLLPSCICAFRGLTTRSTSRASDQ 300
QY 300 TDSP-QRARNHPLHVQVKNLDTGYDPTDDIPAVLGAIDFKGTVFVGASQRDVSGQBOG 358
Db 301 ADTPPRLFNHRWHIQLONLNGTPYDPAEDIPAPLGTDPFRCKVFGVASORNPDS----- 355
QY 359 HYATRAEAHAHIDTTPKYPAPKGLTTLIKSGSDDFNTNQPIRFTPYVGMGDN---WROWEL 415
Db 356 --TTTRAHEAKVDTTSDRFTPKLSLEIITESGDFDNTQSTKFTPYGIVGVDNBAEFQOWSL 413
QY 416 PDYSGRLLTNMLAPAVSPSPGRIILFFRSIVPSAGYGVSGYIDCLIPQEWVQHPYQBA 475
Db 414 PNYSCQFTNNMLAPAVAPNPFCEQLLFFRSQLPSSGSGNSGVLDCLVPQEWVQHPYQBS 473
QY 476 APSQSAVALRVYVNPDTGRTNRIPEAKLHREGFUTVANCNNPIVVPNGYVFRPEAMGNQPY 535
Db 474 APAQTOVALRVYVNPDTGRTNRIPEAKLHKGFMWIAKNGDSPTVPNGYVFRPEWVWNPY 533
QY 536 TLAPNGSGQGRRAQ 550
Db 534 TLAPMGTCGRRRRIQ 548

RESULT 15
Q91725 PRELIMINARY; PRT; 548 AA.
ID Q91725;
AC Q91725;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Towson/313/1994/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171846;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Towson/313/1994/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Towson/313/1994/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
RA Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995.";
RL J. Med. Virol. 53:372-383(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Towson/313/1994/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Towson/313/1994/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414414; AAL12995.1; --
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 59993 MW; 114F3907B3A26D89 CRC64;

```

Query Match		72.8%;	Score 2117;	DB 12;	Length 548;
Best Local Similarity		71.4%;	Pred. No. 1.9e-163;	Indels 12;	Gaps 4;
Matches 396;		Conservative 71;	Mismatches 76;		
Qy	1	MMWASNDAAPSNDGAANLVPEANDEVMAL	EPVVGASIAAPVVGQNIIDPWIRENFVOAP	60	
Db	1	MMWASNDAAPSNDGAAGLVPEINNEAMALE	EPVAGAAIAAPLTGQNIIDPWIMNFVOAP	60	
Qy	61	QGEFTVSPRNSPGEMLLNLELPELNPYL	SHSRMYNGYAGGMQOVVLGNAPFTAGKII	120	
Db	61	QGEFTVSPRNSPGEVLLNLELPEINPYLA	HLARMYNGYAGGFVQVVLGNAPFTAGKVI	120	
Qy	121	FAAVPHFVENISAAQITMCPHVIVDVRQ	LEPVLPLPDINRPFHYNQENTPRMLVA	180	
Db	121	FAAIPNFPIDNLUSAQITMCCHVIVDVRQ	LEPINLPMDVNRNFFHYNQSDSRLRIA	180	
Qy	181	MLYTPLRA-NSGEDVFTVSCRVLTRPAD	PFETFLVPPTVESKTKPFTLILTIGELNS	239	
Db	181	MLYTPLRANSGDDVFTVSCRVLTRPSPD	FSFNLVPPTVESKTKLFTLILTISEMSNS	240	
Qy	240	RFPAAIDMLYTDNESIVVQOPONGRCT	LDGTGTTQTLVPTQICAPRGTLISQTARAADS	299	
Db	241	RFEVPIDSLHTSPENIVVOCNGRVLDG	ELMGTTLQLPSQICAFRGTLTRSTSRASDQ	300	
Qy	300	TDSP-QKARNHPLHVQKNLDGTQYDPTD	DIPAVLGAIDFKGTVFGVASQRDVSGQOEG	358	
Db	301	ADFTPLFNRWHIQLDNLNGTIPYDAED	IAPLGTDPGRKVGFGVASQRNPDS-----	355	
Qy	359	HYATRAHEAHI DTTDPKYAPKLGTLIK	SGDDFNTNQPIRFTPVGMGDN---WROWEL	415	
Db	356	--TTAHEAKVDTTSGRFTPKLASLEI	ITESDDFDNQSTKFTPVGIGVDNESEFQOWSL	413	
Qy	416	PDYSGRLTANNLAPAVSPSPGGRILFF	RSIVPSAGYSGYIDCLIPQEWVQHFYQEA	475	
Db	414	PNYSGQFTHNNLAPAVAPNPLGEQLL	FFRSQLPSSGGRSNGVLDCLVPQEWVQHFYQES	473	
Qy	476	APSQSAVALVRYNPDTRNIFEAKLHRE	GFLTVANGCNPPIVVPNGYERPEAKGNOFY	535	
Db	474	APAQTQVALVRYNPDTRVLFKALHKG	LHGFMTIAKNGDSPITVPPNGYFRFESWNPFI	533	
Qy	536	TLAPMGSGGRRAQ	550		
Db	534	TLAPMGTGRRRIQ	548		

Search completed: June 1, 2004, 13:53:23
Job time : 34.0218 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 46.4752 Seconds
(without alignments)
3289.030 Million cell updates/sec

Title: US-09-926-799-11
Perfect score: 2894
Sequence: 1 MKMASNDAPSDGAGLVP.....VNQFYS LAPVTGKRRRVQ 541

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq_294n04.*
- 1: Geneseqp1980s.*
 - 2: Geneseqp1990s.*
 - 3: Geneseqp2000s.*
 - 4: Geneseqp2001s.*
 - 5: Geneseqp2002s.*
 - 6: Geneseqp2003as.*
 - 7: Geneseqp2003bs.*
 - 8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	2894	100.0	541	4	AAB49710	Aab49710 Small rou
2	1678.5	58.0	548	4	AAB49705	Aab49705 Small rou
3	1656	57.2	535	4	AAB49707	Aab49707 Small rou
4	1650.5	57.0	540	4	AAB49706	Aab49706 Small rou
5	1640.5	56.7	548	5	AAB491272	Aau91272 Norwalk v
6	1613.5	55.8	550	4	AAB49709	Aab49709 Small rou
7	1569	54.2	539	4	AAB49704	Aab49704 Small rou
8	1538.5	53.2	542	4	AAB49708	Aab49708 Small rou
9	1013	35.0	545	4	AAB49700	Aab49700 Small rou
10	1008.5	34.8	530	2	AAR50972	Aar50972 Norwalk v
11	1008.5	34.8	530	7	ADC72176	Adc72176 Norwalk v
12	1005.5	34.7	530	4	AAB49701	Aab49701 Small rou
13	984.5	34.0	530	2	AAR57091	Aar57091 Small rou
14	970	33.5	544	4	AAB49703	Aab49703 Small rou
15	967	33.4	546	4	AAB49702	Aab49702 Small rou
16	309	10.7	579	2	AABW08143	Aaw08143 RHDV caps
17	267.5	9.2	622	4	AAB47045	Aab47045 Feline ca
18	265.5	9.2	623	4	AAB47044	Aab47044 Feline ca
19	265.5	9.1	623	4	AAB47043	Aab47043 Feline ca
20	264	9.1	668	4	AAB67462	Aab67462 Amino aci
21	262.5	9.1	547	4	AAM50108	Aam50108 Feline ca
22	262.5	9.1	671	4	AAM50107	Aam50107 Feline ca
23	254	8.8	668	2	AAR10686	Aar10686 Feline ca
24	254	8.8	668	4	Aae04304	Aae04304 Feline ca
25	250.5	8.7	669	4	AAB67461	Aab67461 Amino aci

26	200	6.9	40	5	AAB49710	Standard; protein; 541 AA.
27	174	6.0	40	5	AAB49710	Standard; protein; 541 AA.
28	119.5	4.1	2209	1	AAP20037	Sequence
29	119	4.1	934	1	AAP20016	Sequence
30	118.5	4.1	1147	5	ABF76724	Foot and
31	117	4.0	2206	2	AAR22210	True type
32	109.5	3.8	1045	7	ADE72434	Human end
33	109.5	3.8	1070	7	ADE72427	Human end
34	109.5	3.8	1093	6	AAO23110	Lig-1 'hu
35	109.5	3.8	1093	7	ADE72425	Human end
36	109.5	3.8	1101	2	AAV08008	Human HLI
37	109.5	3.8	1196	7	ADE72433	Human end
38	108.5	3.7	861	6	ABU38084	Protein e
39	107	3.7	607	5	ABB53295	Human pol
40	106.5	3.7	767	4	AAM25696	Human pro
41	106.5	3.7	861	6	ABF77378	N. gonorr
42	106.5	3.7	861	6	ABU37406	Protein e
43	106.5	3.7	1070	5	ABG69674	Human sec
44	106.5	3.7	1077	5	ABP70116	Human NOV
45	106.5	3.7	1093	5	ABG69660	Human sec

ALIGNMENTS

RESULT 1
AAB49710
ID AAB49710 standard; protein; 541 AA.
XX
AC AAB49710;
XX
DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 11.
XX
KW Small round structured virus; SRSV; food poisoning.
XX
OS Small round structured virus.
XX
PN WO200079280-A1.
XX
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000WO-JP004095.
XX
PR 22-JUN-1999; 99JP-00175928.
XX
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
(DENK-) DENKA SEIKEN KK.
Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
WPI: 2001-080848/09.
N-PSDB; AAF29151.
Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.
Claim 1: Page 64-66; 84pp; Japanese.
This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

Sequence 541 AA;
Query Match 100.0%; Score 2894; DB 4; Length 541;
Best Local Similarity 100.0%; Pred. No. 5e-258;

Wed Jun 2 09:13:25 2004

Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKWASNDAPSDGAGLVEINNEVMPLEPVAGASLATPVVQOQNIIDPWIRNFVQAP 60
 DB 1 MKWASNDAPSDGAGLVEINNEVMPLEPVAGASLATPVVQOQNIIDPWIRNFVQAP 60

QY 61 AGEFTVSPRNSPGEILLDLLELGDPLNPLYLAHARMYNGHAGMEVQIVLAGNATAGKII 120
 DB 61 AGEFTVSPRNSPGEILLDLLELGDPLNPLYLAHARMYNGHAGMEVQIVLAGNATAGKII 120

QY 121 FFAIPGGFYENLSPQITMCPRHVIDVRQLEFFLLPMPDIWNPFHYNQNDPKRLVA 180
 DB 121 FFAIPGGFYENLSPQITMCPRHVIDVRQLEFFLLPMPDIWNPFHYNQNDPKRLVA 180

QY 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDPFEFTFLVPPVESKTKQFALPILKISEMTNS 240
 DB 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDPFEFTFLVPPVESKTKQFALPILKISEMTNS 240

QY 241 RFPVVDVMTARNENOVQVQNGRVTLTGELGTTPLLAIVNICKFKGEVIAKNGDVRSY 300
 DB 241 RFPVVDVMTARNENOVQVQNGRVTLTGELGTTPLLAIVNICKFKGEVIAKNGDVRSY 300

QY 301 RMDMEITNTDGTIDPTEDTGPFGSPDFQILFGVASQNRKNEQNPATRAHEALINTGG 360
 DB 301 RMDMEITNTDGTIDPTEDTGPFGSPDFQILFGVASQNRKNEQNPATRAHEALINTGG 360

QY 361 DHLCPQISSEIYLTSPNIIIRCTNPQLPQSGILGTILIRSDNGHCHDMVGTSPPTPTWP 420
 DB 361 DHLCPQISSEIYLTSPNIIIRCTNPQLPQSGILGTILIRSDNGHCHDMVGTSPPTPTWP 420

QY 421 QQWRRCRSGNCCSSGHRYFPVVMNRVTWIVLSHSGFSTSTRKLPQLNLRWPLIRFIN 480
 DB 421 QQWRRCRSGNCCSSGHRYFPVVMNRVTWIVLSHSGFSTSTRKLPQLNLRWPLIRFIN 480

QY 481 PDTGRVLFEARLHKGFIHTAHTGDNPIVMPNGYFPEAWNQFYS LAPVGTGKRRRV 540
 DB 481 PDTGRVLFEARLHKGFIHTAHTGDNPIVMPNGYFPEAWNQFYS LAPVGTGKRRRV 540

QY 541 Q 541
 DB 541 Q 541

RESULT 2
 AAB49705
 ID AAB49705 standard; protein; 548 AA.

XX AAB49705;
 AC AAB49705;
 DT 04-APR-2001 (first entry)
 XX Small round structured virus protein SEQ ID 6.
 DE Small round structured virus; SRSV; food poisoning.
 XX Small round structured virus.
 OS Small round structured virus.
 XX W0200079280-A1.
 XX 28-DEC-2000.
 XX 22-JUN-2000; 2000WO-JP004095.
 XX 22-JUN-1999; 99JP-00175928.
 XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 XX (DENK-) DENKA SEIKEN KK.
 XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 XX WPI; 2001-080848/09.
 XX N-PSDB; AAF29146.

PT Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 XX Claim 1; Page 52-54; 84pp; Japanese.
 XX This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX

Sequence 548 AA;

Query Match 58.0%; Score 1678.5; DB 4; Length 548;
 Best Local Similarity 58.9%; Pred. No. 9.9e-146; Indels 45; Gaps 7;
 Matches 334; Conservative 67; Mismatches 121;

QY 1 MKWASNDAPSDGAGLVEINNEVMPLEPVAGASLATPVVQOQNIIDPWIRNFVQAP 60
 DB 1 MKWASNDAPSDGAGLVEINNEVMPLEPVAGASLATPVVQOQNIIDPWIRNFVQAP 60

QY 61 AGEFTVSPRNSPGEILLDLLELGDPLNPLYLAHARMYNGHAGMEVQIVLAGNATAGKII 120
 DB 61 AGEFTVSPRNSPGEILLDLLELGDPLNPLYLAHARMYNGHAGMEVQIVLAGNATAGKII 120

QY 121 FFAIPGGFYENLSPQITMCPRHVIDVRQLEFFLLPMPDIWNPFHYNQNDPKRLVA 180
 DB 121 FFAIPGGFYENLSPQITMCPRHVIDVRQLEFFLLPMPDIWNPFHYNQNDPKRLVA 180

QY 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDPFEFTFLVPPVESKTKQFALPILKISEMTNS 240
 DB 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDPFEFTFLVPPVESKTKQFALPILKISEMTNS 240

QY 241 RFPVVDVMTARNENOVQVQNGRVTLTGELGTTPLLAIVNICKFKGEVIAKNGDVRS - 299
 DB 241 RFPVVDVMTARNENOVQVQNGRVTLTGELGTTPLLAIVNICKFKGEVIAKNGDVRS - 299

QY 300 -----YRMDMEITNTDGTIDPTEDTGPFGSPDFQILFGVASQNRKNEQNP 348
 DB 300 QADTATPRLFNYWHVQDLNLTNGTTPYDPAEDIPGLGTDPFRGVGVASQNRNLD -T 356

QY 349 TRAHEAIIITGGDHLCPQISSEIYLTSPNIIIRCTNPQLPQSGILGTILIRSDN - 403
 DB 349 TRAHEAIIITGGDHLCPQISSEIYLTSPNIIIRCTNPQLPQSGILGTILIRSDN - 403

QY 357 TRAHEAKVDITTAGRFTPKLGSLEISTDSDDDPDQ-NQPTKFTPVG-----IGVDNEAEFQ 409
 DB 357 TRAHEAKVDITTAGRFTPKLGSLEISTDSDDDPDQ-NQPTKFTPVG-----IGVDNEAEFQ 409

QY 404 -----GHCHDMVGTSPPTTTPWQWRRCRSGNCCSSGHRYFPVVMNRVTWIVLS 454
 DB 404 -----GHCHDMVGTSPPTTTPWQWRRCRSGNCCSSGHRYFPVVMNRVTWIVLS 454

QY 410 QWSLDPDYSQGTTHNWNLAFAVAFNFPFBQLLFFRSQPSGSGRNGVLDCLVPOEWOHF 469
 DB 410 QWSLDPDYSQGTTHNWNLAFAVAFNFPFBQLLFFRSQPSGSGRNGVLDCLVPOEWOHF 469

QY 455 HKSFGSTSTRKLPQLNLRWPLIRFINPDTGRVLFEARLHKGFIHTAHTGDNPIVMPNG 514
 DB 455 HKSFGSTSTRKLPQLNLRWPLIRFINPDTGRVLFEARLHKGFIHTAHTGDNPIVMPNG 514

QY 470 YQESAPAQOV-----ALVRYNPDPTGKVLFEAKLHKLGFMTIANGDSPTVPNG 521
 DB 470 YQESAPAQOV-----ALVRYNPDPTGKVLFEAKLHKLGFMTIANGDSPTVPNG 521

QY 515 YFRFEAWNQFYS LAPVGTGKRRVQ 541
 DB 522 YFRFEAWNQFYS LAPVGTGKRRVQ 548

RESULT 3
 AAB49707
 ID AAB49707 standard; protein; 535 AA.
 XX AAB49707;
 AC AAB49707;
 DT 04-APR-2001 (first entry)
 XX Small round structured virus protein SEQ ID 8.
 DE Small round structured virus; SRSV; food poisoning.
 XX Small round structured virus; SRSV; food poisoning.

OS Small round structured virus.
PN WO200079280-A1.
XX 28-DEC-2000.
XX 22-JUN-2000; 2000WO-JP004095.
XX 22-JUN-1999; 99JP-00175928.
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI; 2001-080848/09.
XX N-PSDB; AAF29148.
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX strains for investigation of food poisoning outbreaks, contains
XX antibodies.
XX Claim 1; Page 57-59; 84pp; Japanese.
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks
XX Sequence 535 AA;
Query Match 57.2%; Score 1656; DB 4; Length 535;
Best Local Similarity 59.2%; Pred. No. 1.1e-143;
Matches 330; Conservative 67; Mismatches 122; Indels 38; Gaps 7;
QY 1 MKWASNDAAPSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFVQAP 60
DB 1 MKWASNDAAPSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFVQAP 60
QY 61 AGEFTVSPRNSGCEILLDLGLPDLPNLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120
DB 61 AGEFTVSPRNSGCEILLDLGLPDLPNLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120
QY 121 FFAIPPGPYENLSPSQITMCPHVLIIDVROLEPFLPMDIWNFFHYNQGNPKRLVA 180
DB 121 FFAIPPGPYENLSPSQITMCPHVLIIDVROLEPFLPMDIWNFFHYNQGNPKRLVA 180
QY 181 MLYTPLRANSGDDVFTVSCRVLTKPSDFEFTFLVPPTVESKTKQFALPILKISEMTNS 240
DB 181 MLYTPLRANSGDDVFTVSCRVLTKPSDFEFTFLVPPTVESKTKQFALPILKISEMTNS 240
QY 241 RPPVPVDMYTARNENQVVPQNGRVCALDGEQGTQLLPTAICFRGRINQKVSQE--N 298
DB 241 RPPVPVDMYTARNENQVVPQNGRVCALDGEQGTQLLPTAICFRGRINQKVSQE--N 298
QY 300 YEMDEITNTDCTPDIPTEDTPGICSPDFQGLFVASQRNKNQNPATRAHEAIINTG 359
DB 299 HWNNQVNTINGTPDPTGDVPAPLGTDPFSKLFVLSQR---DHDNACRSHDAVIATN 355
QY 360 GDHLCPOISSSIYLTSPNLRCTNPQPLPQSGL-----RGTLIRSDNGH 405
DB 356 SAKFTPKLGAIGTWEEDDVHINQTKFTPVGLFENEGFNQWTLPNYSGALTL----- 409
QY 406 CHDMVGTSTPTTWPQQWRRCRSGNCCSSGHRYPVVVMNRVTIV-LSHKSGFSTSTR 464
DB 410 --NMGLAPPVAPTFPGEQLFRSHIPLKGGVADPVIDCLLPQEWIQLYQSAPSQSD- 466
QY 465 KLPQLNLRLPRLFRINPDGTGRVLFEARLHKQCFITVANTGNDPNIYMPNGYFRPRAWNO 524
DB 467 -----VALIRFTNPDTGRVLFEARLHRSYITVANTGSRPIVVPANGYFRFTWVNO 518

QY 525 FYSLAPVGTGKRRRVQ 541
DB 519 FYSLAPMGTGNGRRVQ 535
RESULT 4
AAB49706
ID AAB49706 standard; protein; 540 AA.
XX AAB49706;
XX 04-APR-2001 (first entry)
XX Small round structured virus protein SEQ ID 7.
XX Small round structured virus; SRSV; food poisoning.
XX Small round structured virus.
XX WO200079280-A1.
XX 28-DEC-2000.
XX 22-JUN-2000; 2000WO-JP004095.
XX 22-JUN-1999; 99JP-00175928.
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI; 2001-080848/09.
XX N-PSDB; AAF29147.
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX strains for investigation of food poisoning outbreaks, contains
XX antibodies.
XX Claim 1; Page 54-57; 84pp; Japanese.
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks
XX Sequence 540 AA;
Query Match 57.0%; Score 1650.5; DB 4; Length 540;
Best Local Similarity 56.5%; Pred. No. 3.7e-143;
Matches 330; Conservative 65; Mismatches 102; Indels 87; Gaps 8;
QY 1 MKWASNDAAPSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFVQAP 60
DB 1 MKWASNDATPSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFVQAP 60
QY 61 AGEFTVSPRNSGCEILLDLGLPDLPNLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120
DB 61 AGEFTVSPRNSGCEILLDLGLPDLPNLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120
QY 121 FFAIPPGPYENLSPSQITMCPHVLIIDVROLEPFLPMDIWNFFHYNQGNPKRLVA 180
DB 121 FFAIPPGPYENLSPSQITMCPHVLIIDVROLEPFLPMDIWNFFHYNQGNPKRLVA 180
QY 181 MLYTPLRANSGDDVFTVSCRVLTKPSDFEFTFLVPPTVESKTKQFALPILKISEMTNS 240
DB 181 MLYTPLRANSGDDVFTVSCRVLTKPSDFEFTFLVPPTVESKTKQFALPILKISEMTNS 240
QY 241 RPPVPVDMYTARNENQVVPQNGRVCALDGEQGTQLLPTAICFRGRINQKVSQE--N 299

Db	241	RPFLSIDEMVTPSNESIVVQNGRVTLDGELLGTTQLQACNICSIIRKVV---TGQVPSE	297
Qy	300	YRDMETITNDGTDPIDTDPGIGSPDFQGLFGVASORNKNEQNCPATRAHEALINT	358
Db	298	QHMMLEITNLNGTQDFTDDVPAPLGVDPDFAGEVFGVLSORNGESPNARAHAVAT	357
Qy	359	GDHLCPOISSSEIYLTSPNILRCTNPQLPQSLGRGLTILIRSDNGHCHDMVGTSPTPT	418
Db	358	YSDKTKPKLGLVQI-----GWTNTNDV	379
Qy	419	WPQWRRCRSGNCCSSGHRV-----PVFVMMRVTVIV-----LS	454
Db	380	ENQTKFTPIGLNEVANGHRPEQWTLPRYSGALTLMMNLAPAVLFPGERLLFFRSKVP	439
Qy	455	HKSGFSTST--RKLQOLNRW-----PLIRFINPDTGRVLFEARLHKQGF	497
Db	440	LRGGFGNPAIDCSVPQ---EWQHFQESAPSGLGDVALVRYVNPDTGRVLFEARLHKQGF	496
Qy	498	ITVAHTGDNPIVMPNGYFRFEAWNQFYSLAPVGTGKRRRVQ	541
Db	497	LTVSSSTGCVVVPVANGYFKFDSWVNFYSLAPMGTGNGRRRVQ	540
RESULT 5			
ID	AAU91272	standard; protein; 548 AA.	
XX	AAU91272;		
DT	18-JUN-2002	(first entry)	
DE	Norwalk virus associated polynucleotide #1.		
XX	Nowalk virus; monoclonal antibody; geno group I; geno group II;		
KW	immunological detection; food; viral infection.		
OS	Norwalk virus.		
XX	JP2002020399-A.		
PN	23-JAN-2002.		
XX	10-JUL-2000; 2000JP-00208151.		
PF	10-JUL-2000; 2000JP-00208151.		
PR	(OSAP) OSAKA PREFECTURE.		
PA	(NISE-) NIPPON SEIBUTSU KAGAKU CENT KK.		
PA	(IATR) IATRON LAB INC.		
XX	WPI; 2002-287412/33.		
DR	A monoclonal antibody useful in the immunological detection and diagnosis		
XX	of Norwalk virus infection.		
PT	Disclosure; Page 12-13; 24pp; Japanese.		
PS	The invention describes a monoclonal antibody recognising Norwalk virus,		
XX	a capsid protein of Norwalk virus, or a common antigen epitope on the		
CC	capsid protein molecule of geno group I and geno group II. The antibody		
CC	is useful for immunological detection and quantitative analysis of		
CC	Norwalk virus in foods and the serum of infected patients. This sequence		
CC	represents a Norwalk virus associated protein described in the invention		
XX	Sequence 548 AA;		
Qy	Query Match	56.7%; Score 1640.5; DB 5; Length 548;	
XX	Best Local Similarity	58.6%; Pred. No. 3.2e-142;	
XX	Matches 330; Conservative	66; Mismatches 122; Indels 45; Gaps 8;	
Qy	1	MRMASNDAAAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQONIIDPIMMNFVOAP	60

PS Claim 1; Page 62-64; 84pp; Japanese.

XX This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX Sequence 550 AA;

Query Match 55.8%; Score 1613.5; DB 4; Length 550;
 Best Local Similarity 55.3%; Pred. No. 1e-139; Mismatches 99; Gaps 12;
 Matches 329; Conservative 77;

QY 1 MKMASNDAAAPSDGAGLVPEINNEVMPLEPVAGASLATPVGQQNIIDPWIRNPFVQAP 60
 DB 1 MKMASNDAAAPSDGAGLVPEINNEVMPLEPVAGASLATPVGQQNIIDPWIRNPFVQAP 60

QY 61 AGEFTVSPRNSPGEILLDLLELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120
 DB 61 QGEFTVSPRNSPGEILLDLLELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120

QY 121 FPAIPGPPYENLSQITMCHVVIDVROLEPFLPMPDINWNNFFHYNQGNPKLRLVA 180
 DB 121 FPAIPGPPYENLSQITMCHVVIDVROLEPFLPMPDINWNNFFHYNQGNPKLRLVA 180

QY 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDFEFLVPPTVESKTKQFALPILKISEMTNS 240
 DB 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDFEFLVPPTVESKTKQFALPILKISEMTNS 240

QY 241 RFPPVDMYTARNENQVVPQNGRVTLDGELLGTTPLLAVNICKPKGEVIARNGDVRSY 300
 DB 241 RFPPVDMYTARNENQVVPQNGRVTLDGELLGTTPLLAVNICKPKGEVIARNGDVRSY 300

QY 301 RMDMEITNDGTPIODPTEDTPGIPGSPDFQIGILFGVASQNKQNPATRAHEAINTGG 360
 DB 301 RMDMEITNDGTPIODPTEDTPGIPGSPDFQIGILFGVASQNKQNPATRAHEAINTGG 360

QY 487 LFPEARLHKOGFTVAHTGDNPIVMPNGYFRFEAMVNOFYSLAPVGTGKGRRRVQ 541
 DB 487 LFPEARLHKOGFTVAHTGDNPIVMPNGYFRFEAMVNOFYSLAPVGTGKGRRRVQ 541

QY 496 IFEAKLHREGFLTVANGCNPIVPPNGYFRFEAMGNQFYTLAPMGSGGRRRAQ 550
 DB 496 IFEAKLHREGFLTVANGCNPIVPPNGYFRFEAMGNQFYTLAPMGSGGRRRAQ 550

RESULT 7
 AAB49704
 ID AAB49704 standard; protein; 539 AA.
 XX AAB49704;
 AC AAB49704;
 XX AAB49704;
 DT 04-APR-2001 (first entry)
 XX Small round structured virus protein SEQ ID 5.
 DE Small round structured virus; SRSV; food poisoning.
 XX Small round structured virus.
 OS Small round structured virus.
 XX WO200079280-A1.
 PN WO200079280-A1.
 XX

PD 28-DEC-2000.

XX 22-JUN-2000; 2000WO-JP004095.
 XX 22-JUN-1999; 99JP-00175928.
 PR (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.
 XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 PI WPI: 2001-080848/09.
 XX N-PSDB; AAF29145.
 DR Kit for the detection and typing of small round-structured virus (SRSV)
 XX strains for investigation of food poisoning outbreaks, contains
 PT antibodies.
 PT Claim 1; Page 50-52; 84pp; Japanese.
 PS This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks
 XX Sequence 539 AA;

Query Match 54.2%; Score 1569; DB 4; Length 539;
 Best Local Similarity 54.8%; Pred. No. 1.3e-135;
 Matches 308; Conservative 76; Mismatches 130; Indels 48; Gaps 9;

QY 1 MKMASNDAAAPSDGAGLVPEINNEVMPLEPVAGASLATPVGQQNIIDPWIRNPFVQAP 60
 DB 1 MKMASNDAAAPSDGAGLVPEINNEVMPLEPVAGASLATPVGQQNIIDPWIRNPFVQAP 60

QY 61 AGEFTVSPRNSPGEILLDLLELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120
 DB 61 AGEFTVSPRNSPGEILLDLLELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120

QY 121 FPAIPGPPYENLSQITMCHVVIDVROLEPFLPMPDINWNNFFHYNQGNPKLRLVA 180
 DB 121 FPAIPGPPYENLSQITMCHVVIDVROLEPFLPMPDINWNNFFHYNQGNPKLRLVA 180

QY 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDFEFLVPPTVESKTKQFALPILKISEMTNS 240
 DB 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDFEFLVPPTVESKTKQFALPILKISEMTNS 240

QY 241 RFPPVDMYTARNENQVVPQNGRVTLDGELLGTTPLLAVNICKPKGEVIARNGDVRSY 300
 DB 241 RFPPVDMYTARNENQVVPQNGRVTLDGELLGTTPLLAVNICKPKGEVIARNGDVRSY 300

QY 301 RMDMEITNDGTPIODPTEDTPGIPGSPDFQIGILFGVASQNKQNPATRAHEAINTGG 360
 DB 301 RMDMEITNDGTPIODPTEDTPGIPGSPDFQIGILFGVASQNKQNPATRAHEAINTGG 360

QY 456 WYQHFCEBAAPQAQSDVALLRFVNPDTGRVLFCKLHKGYSYVAHTGPHDLVIPNGYPR 515
 DB 456 WYQHFCEBAAPQAQSDVALLRFVNPDTGRVLFCKLHKGYSYVAHTGPHDLVIPNGYPR 515

QY 518 FEAWNQYSLAPVGTGKGRRR 539
 DB 518 FEAWNQYSLAPVGTGKGRRR 539

QY 516 FDSWYNQFYTLAPMGNGAGRRR 537
 DB 516 FDSWYNQFYTLAPMGNGAGRRR 537

Db 301 NL-LQTYPNGASYDPTDEVPAFLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGIYSTT 359
Qy 360 GDHLCPOISSSEIYLTSPNILECTNPQPLQSGLRGTILIRSDNGHCHDMV----- 410
Db 360 SGKFTPKIGSIGLHSITEHV-----HPNQSRFTTPGVAVDENTPFOQWVLPHYAGSLA 413
Qy 411 ---GTSP-TTPTWPOQWRRCSRSGNSCCSGHRYPPVVMNRVTWVLSHSGSFSTSKL 466
Db 414 LNTNLAPAVAPTFPGEQLLFFRSRVPVCG-----LQQDAFIDCL--L 455
Qy 467 POLNLRW-----PLIRFNPDTGRVLFEARLHKOGFITVAHTGDNPIVMP 511
Db 456 PQ---EWNHIFYQEAAPSQADVALLIRYNPDTRTLFEAKLHRSFGFITVSHGTGALPLVVP 512
Qy 512 PNGYFRFEAWNQFYSLAPVGTGKGRRYQ 541
Db 513 PNGHFRFDSWVNOFYSLAPMGTGNGRRRIQ 542

RESULT 9
AAB49700 standard; protein; 545 AA.

XX AAB49700;
XX 04-APR-2001 (first entry)
XX Small round structured virus protein SEQ ID 1.
XX Small round structured virus; SRSV; food poisoning.
XX Small round structured virus.
XX WO200079280-A1.
XX 28-DEC-2000.
XX 22-JUN-2000; 2000WO-JP004095.
XX 22-JUN-1999; 99JP-00175928.
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI; 2001-080848/09.
XX N-PSDB; AAF29141.
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX strains for investigation of food poisoning outbreaks, contains
XX antibodies.
XX Claim 1; Page 40-42; 84pp; Japanese.
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks

Query Match 35.0%; Score 1013; DB 4; Length 545;
Best Local Similarity 40.5%; Pred. No. 3e-84;
Matches 240; Conservative 79; Mismatches 154; Indels 120; Gaps 18;
Qy 1 MKMASNDAAAPSSDGAAG---LVPEINN-EVMPLEPVAGASLATPVPVQGNQIIDPWIRNF 56
Db 1 MMWASKDAPTMDGTSGAGQLVPEANTAEPISEPVAGATAATAATAGVNMIDPWIMNY 60

RESULT 8
ID AAB49708 standard; protein; 542 AA.
XX AAB49708;
XX 04-APR-2001 (first entry)
XX Small round structured virus protein SEQ ID 9.
XX Small round structured virus; SRSV; food poisoning.
XX Small round structured virus.
XX WO200079280-A1.
XX 28-DEC-2000.
XX 22-JUN-2000; 2000WO-JP004095.
XX 22-JUN-1999; 99JP-00175928.
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI; 2001-080848/09.
XX N-PSDB; AAF29149.
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX strains for investigation of food poisoning outbreaks, contains
XX antibodies.
XX Claim 1; Page 59-61; 84pp; Japanese.
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks

Query Match 53.2%; Score 1538.5; DB 4; Length 542;
Best Local Similarity 54.4%; Pred. No. 8.4e-133; Indels 57; Gaps 9;
Matches 310; Conservative 72; Mismatches 131;
Qy 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVPVQGNQIIDPWIRNFVQAP 60
Db 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVPVQGNQIIDPWIRNFVQAP 60
Qy 61 AGEFTVSPNSPGEILLDLLEQPDLPNLYLAHRYNGHAGMEQVIVLAGNAFTAGKII 120
Db 61 NGEFTVSPNSPGEILLDLLEQPDLPNLYLAHRYNGHAGMEQVIVLAGNAFTAGKIL 120
Qy 121 FRAIPGFPVENLSPQITMCHPHVIDVRQLEPFLPMPDIWNFFHYNQGNPKLRJVA 180
Db 121 FRAIPGFPVENLSPQITMCHPHVIDVRQLEPFLPMPDIWNFFHYNQGNPKLRJVA 180
Qy 181 MLYTPLRANSGDDVFTVSCRVLTKPSDFETFLVPPVESKTKQFALPKIKSEMTNS 240
Db 181 MLYTPLRANSGDDVFTVSCRVLTKPSDFETFLVPPVESKTKQFALPKIKSEMTNS 240
Qy 241 RFPVPDVMYATARNQVQPNQGRVTLDELGLLTPPLAVNICPKFGKGVIAKNGDVRSY 300
Db 241 RFPPIEQLYATARNQVQPNQGRVTLDELGLLTPPLAVNICPKFGKGVIAKNGDVRSY 300
Qy 301 RMDMEITNDGTIDPTEDTFFGPIGSPDFGILFGVASQRNKN-EQNPAATRAHEALINTG 359

QY 57 VQAPAGEFTVSPRNSPGEILLDLELGPDLNPLYLAHARMYNGHAGMEVOIVLAGNAFTA 116
 DB 61 VQAPAGEFTISPNTTGDILFDLQGLPHLNPFLSHLAQYNGWGVNKMVKVLAGNAFTA 120
 QY 117 GKIIIPAAIPGPPYENLSPQITMCPHVIDVRLQLEPFLPMPDIWNNEFFHYNOGNDPKL 176
 DB 121 GKIIISCIPGFAAQNISIAQATMFHVIADVRVLEPFLDVEDVRNVLPH--NNDNAPTM 179
 QY 177 RLVMALYTLPLRA--NNSGDDVFTVSCRVLTKPSDFEFTLVPPPTVESKTKQFALPILKI 234
 DB 180 RLVMALYTLPLRASGSSGTDPPFIAGRVLTCPSPDFSLFLVPPNVEQTKPFSVENLPL 239
 QY 235 SEMTNSRFPVVDVMTARNENQVOPQNGRVTLDELIGLTTPLLAVINCKFKGEVIAQN 294
 DB 240 NTLNSRVPFLIKSMVMVRDHGMQVQFQNGRVTLQGLQGTTPTSASQLCKIRGSVFHAN 299
 QY 295 GQVRSVRMDMEINTDGTPTDPTEDTGPFGIPGSDFGQILFGVASQRNKNQEQNPATRAHEA 354
 DB 300 GG-NGY----NLTELDGSPYHAF--SPADIGPDLGECDDHM-----EASFTTQ---- 342
 QY 355 IINTGGDLCPQISSEIYLTSPNLRCTNPQPLQSGRLGTILI--RSDNGHCHDMVG- 411
 DB 343 -FNTG--DIVIKQINVKQESAFAPHL-----GFIQADGLSDVSNTNMIK 384
 QY 412 ----TSPTTTPWQWRRCRSGNCCSGHRYPVVVMNRVTWIVLSHKSGFSTSTRKLPQ 468
 DB 385 LGWVSPV-----SDGHRGDVD-----PWVPIPRYGSTLTTEAAQLAPP 420
 QY 469 L-----NLWLP-----LIRFINDP 482
 DB 421 IYPPGGEIAIVFMSDFPIAHGTNGLSVPCITPQEFVTHFVNEQAPTRGEAALLHYLDDP 480
 QY 483 TGRVLFPEARLHKQGFITVA--HTGDNPIVMPNGYERFEAWNOFYSLAPVGT 533
 DB 481 THRNLFEPKLYEGFWTCVPSNGSGTGPQLPFGVGVFVSWRSFYQLKPVGT 533

RESULT 10
 AAR50972
 ID AAR50972 standard; protein; 530 AA.
 XX AC AAR50972;
 XX DT 16-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 05-OCT-1994 (first entry)
 XX DE Norwalk virus strain 8FIIa protein (encoded by ORF2).
 XX KW Norwalk virus; pathogen; acute gastroenteritis; food poisoning;
 KW seafood contamination; diagnostic assay; calcivirus; small round virus.
 XX OS Norwalk virus; (strain 8FIIa).
 XX PN W09405700-A2.
 XX PD 17-MAR-1994.
 XX PF 07-SEP-1993; 93WO-US008447.
 XX PR 07-SEP-1992; 92US-00941365.
 XX PA (BAYU) BAYLOR COLLEGE MEDICINE.
 XX PI Matson DO, Estes MK, Jiang X, Graham DY;
 XX WPI; 1994-101125/12.
 XX N-PSDB; AAQ56826.
 XX PT DNA from Norwalk and related viruses - used for preparing prods. for use
 PT in diagnostic assays, detection and vaccines for Norwalk and related
 PT viruses.

XX Claim 14; Page 68-70; 156pp; English.
 XX The Norwalk virus was isolated from stool samples from adult volunteers
 CC infected with safety tested Norwalk virus strain 8FIIa. The coding
 CC sequence is useful for the design of probes for use in diagnostic assays
 CC for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct
 CC FN field.) (Updated on 16-OCT-2003 to standardise OS field)
 XX SQ Sequence 530 AA;
 Query Match 34.8%; Score 1008.5; DB 2; Length 530;
 Best Local Similarity 40.9%; Pred. No. 7.4e-84;
 Matches 235; Conservative 72; Mismatches 180; Indels 87; Gaps 14;
 QY 1 MKVASNDAAAPSSDGAAG---LVPEIN-NEVMPLBEPVAGASLATPVVGOQNIIDPWRNPF 56
 DB 1 MMWASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGVQNPIDPWLNPF 60
 QY 57 VQAPAGEFTVSPRNSPGEILLDLELGPDLNPLYLAHARMYNGHAGMEVOIVLAGNAFTA 116
 DB 61 VQAPAGEFTISPNTTGDILFDLQGLPHLNPFLSHLAQYNGWGVNKMVKVLAGNAFTA 120
 QY 117 GKIIIPAAIPGPPYENLSPQITMCPHVIDVRLQLEPFLPMPDIWNNEFFHYNOGNDPKL 176
 DB 121 GKIIISCIPGFAAQNISIAQATMFHVIADVRVLEPFLDVEDVRNVLPHNDRNQQT 180
 QY 177 RLVMALYTLPLRASGSSGTDPPFIAGRVLTCPSPDFSLFLVPPNVEQTKPFSVENLPL 236
 DB 180 RLVMALYTLPLRASGSSGTDPPFIAGRVLTCPSPDFSLFLVPPNVEQTKPFSVENLPL 240
 QY 237 MTNSRFPVVDVMTARNENQVOPQNGRVTLDELIGLTTPLLAVINCKFKGEVIAKNGD 296
 DB 241 LNSRAPLPISSMGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRG---TSNGT 297
 QY 297 VRSYRMDMEINTDGTPTDPTEDTGPFGIPGSDFGQILFGVASQRNKNQEQNPAT 349
 DB 298 V-----INLTEDGTTPHFFEG--PAPIGFPLGGLGCDHNNWTQGHSSQTYD----- 344
 QY 350 RAHEAIINTGGDLCPQISSEIYLTSPNLRCTNPQPLQSGRLGTILIIRSDNGHCHDM 409
 DB 345 -----VDTPPTFPVHLGS-----IQANGISGNY 369
 QY 410 VG-----TSPTTTPWQ--QMRRCRSGNCCSGHRYP--VPVVMNRVTWIVLSHKSGFS 460
 DB 370 VGVLSWISPPSPHSGSQVDLWKIPNYGSSITEATHLAPSVYPPGFGVLFVFMKMPGPG 429
 QY 461 TSTRK--LPQLNL-----RWPLIRPINDTGRVLFPEARLHKQGFITVAHTG-- 504
 DB 430 AYNLPCLLPQEIYISHLASEQAPTVGEAALLHYVDPDTCRNLGFEKAYPDGFLTCVPNGAS 489
 QY 505 DNPIVMPNGYERFEAWNOFYSLAPVGTGKRR 538
 DB 490 SGFQQLPINGVFVFSWRSFYQLKPVGTASSAR 523

RESULT 11
 ADC72176
 ID ADC72176 standard; protein; 530 AA.
 XX AC ADC72176;
 XX DT 18-DEC-2003 (first entry)
 XX DE Norwalk virus protein 2 amino acid sequence.
 XX KW immune response; non-Norwalk virus agent; immunogen; Norwalk virus;
 KW viral pathogen; acute gastroenteritis; virucidal; antiviral vaccine.
 XX OS Norwalk virus.
 XX PN US6572862-B1.
 XX

Wed Jun 2 09:13:25 2004

PD 03-JUN-2003.
XX 07-JUN-1995; 95US-00486049.
XX 08-NOV-1989; 89US-00433492.
XX 27-APR-1990; 90US-00515993.
PR 27-AUG-1990; 90US-00573509.
PR 06-MAY-1991; 91US-00696454.
XX (BAYU) BAYLOR COLLEGE MEDICINE.
PA
XX
PI Estes MK, Jiang X, Graham DY;
XX
XX WPI; 2003-776005/73.
DR N-PSDB; ADC72174.
XX
XX Inducing an immune response against non-Norwalk virus agents, comprises
PT administering an immunogen recombinantly expressed from a cDNA from
PT Norwalk virus.
XX
XX
PS Example 4; SEQ ID NO 3; 45pp; English.
XX
XX This invention relates to a novel method of inducing an immune response
CC in an individual against Norwalk virus and non-Norwalk virus agents, by
CC orally or parenterally administering an immunogen recombinantly expressed
CC or synthesised from a cDNA of Norwalk virus given in the specification.
CC Norwalk virus is one of the most important viral pathogens, causing acute
CC gastroenteritis. The invention may be used for the development of
CC compounds with virucidal activity or an antiviral vaccine. The present
CC sequence is the amino acid sequence of a protein encoded by the Norwalk
CC virus genome of the invention.
XX
XX
SQ Sequence 530 AA;
Query Match 34.8%; Score 1008.5; DB 7; Length 530;
Best Local Similarity 40.9%; Pred. No. 7.4e-84;
Matches 235; Conservative 72; Mismatches 180; Indels 87; Gaps 14;
QY 1 MKMASNDAAAPSSDGAAG---LVPEIN-NEVMPLFPVAGASLAPVVGQNIIDPWRNNF 56
DB 1 MMWASKDATSSVDGASGAGQLVPEVNASDPLAMDPLAVGASTAVATAGQVNPIDPWIINN 60
QY 57 VOAPAGEFTVSPRNSPGEILLDLGLPDLNPLYLAHARMYNGHAGMEVOIVLAGNAFTA 116
DB 61 VOAPQGEFTISPNNTPGGVLFDSLGLPHLNFLLHLSQMYNGWGNVRIMLAGNAFTA 120
QY 117 GKIIIFAAIIPPGPFYENLSPSQITMCPHVIIDVROLEPFLPMPDIWNNFHYNQNDPKL 176
DB 121 GKIIIVSCIPPGFGSHNLITIAQATLFPHVIADVRLDPIEVPLEDVRNVLFNHNDRNQOTM 180
QY 177 RLVMALYTPLRANSGDDVFTVSCRLTKPSDREFTFLVPPTVESKTKQFALPKIKISE 236
DB 181 RLVCMLYTPLRGTGGGTGDSFVAGRWMTCPSPDFNPLFLVPPTVEQKTRPTLPLPLSS 240
QY 237 MTNSRFPVPVDMYTARNENQVQPNQGRVTLGDBELLGTTPLLANVICFKGEVIKNGD 296
DB 241 LNSRAPLFISSMGISPDNVQVQFQNGRCTLGRLVGTTPVSLSHVAKIRG---TSNGT 297
QY 297 VRSYRMDMEITNDGTPIPTEDTPGICSPDFQG-----ILFGVASQRKNQONPAT 349
DB 298 V-----INLTLDGTFPHFPEG-PAPIGPDLGGCDWHINNTQFGHSQTOYD----- 344
QY 350 RAHEALINTGDHLCQISSEIYLTSPNILRCTNPQLPQSLRGTILIRSDNGCHDM 409
DB 345 -----VDITPDTFVPHLGS-----IQANGIGSGNY 369
QY 410 VG-----TSPTTPTWFO--QWRRCRSGNCCSGCHRYP--VPVMNRVTWIVLSHKSGPS 460
DB 370 VGVLSMTISPSPSHSGSQVDLWKIPNYGSSITEATHLAPSYPPGFGFVLFVFFWMSKMPGEG 429
QY 461 TSTRK--LPQLN-----RWPLIRFINTDTRGVLFPEARLHKQGFITVAHTG-- 504
DB 430 AYNLPCLLPQEIYSHLASEQAPTVEGAALLHYVDPTGRNLGFEKAYPDGFLTCVPNGAS 489

QY 505 DNPIVMPNNGYFRPEAWVNFQFYSLAPVGTGKRR 538
DB 490 SGFQQLPINGVFVSVWSRFYQLKPVGTASSAR 523
RESULT 12
AAB49701
ID AAB49701 standard; protein; 530 AA.
XX
XX AAB49701;
XX AC
XX 04-APR-2001 (first entry)
XX
XX Small round structured virus protein SEQ ID 2.
XX
XX Small round structured virus; SRSV; food poisoning.
XX
XX Small round structured virus.
XX
XX WO200079280-A1.
XX
XX 28-DEC-2000.
XX
XX 22-JUN-2000; 2000WO-JP004095.
XX
XX 22-JUN-1999; 99JP-00175928.
XX
XX (NINA-) JAPAN NAT INST INSPECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI; 2001-080848/09.
XX N-PSDB; AAF29142.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
XX Claim 1; Page 42-45; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710, -
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX
XX
SQ Sequence 530 AA;
Query Match 34.7%; Score 1005.5; DB 4; Length 530;
Best Local Similarity 41.0%; Pred. No. 1.4e-83;
Matches 236; Conservative 70; Mismatches 179; Indels 91; Gaps 15;
QY 1 MKMASNDAAAPSSDGAAG---LVPEIN-NEVMPLFPVAGASLAPVVGQNIIDPWRNNF 56
DB 1 MMWASKDATSSVDGASGAGQLVPEVNASDPLAMDPLAVGASTAVATAGQVNPIDPWIINN 60
QY 57 VOAPAGEFTVSPRNSPGEILLDLGLPDLNPLYLAHARMYNGHAGMEVOIVLAGNAFTA 116
DB 61 VOAPQGEFTISPNNTPGGVLFDSLGLPHLNFLLHLSQMYNGWGNVRIMLAGNAFTA 120
QY 117 GKIIIFAAIIPPGPFYENLSPSQITMCPHVIIDVROLEPFLPMPDIWNNFHYNQNDPKL 176
DB 121 GKIIIVSCIPPGFGSHNLITIAQATLFPHVIADVRLDPIEVPLEDVRNVLFNHNDRNQOTM 180
QY 177 RLVMALYTPLRANSGDDVFTVSCRLTKPSDREFTFLVPPTVESKTKQFALPKIKISE 236
DB 181 RLVCMLYTPLRGTGGGTGDSFVAGRWMTCPSPDFNPLFLVPPTVEQKTRPTLPLPLSS 240
QY 237 MTNSRFPVPVDMYTARNENQVQPNQGRVTLGDBELLGTTPLLANVICFKGEVIKNGD 296

241	LSN	GRAP	PI	SGM	GIS	PN	VS	QF	QNG	RCT	LD	GR	LV	GT	TP	VS	LS	HV	AK	IR	---	TS	NGT	297
297	VS	RY	MD	ME	IT	NT	DT	DP	TD	ET	DP	GI	CS	PD	FQ	---	IL	FG	VS	AQ	RN	KE	ON	349
298	V	---	IN	LT	EL	DT	PH	PE	FG	---	PA	PI	GP	DL	GC	DW	HI	NT	QF	GH	SS	QT	YD	344
350	RA	EA	IN	TG	GH	LC	PO	IS	---	SE	YT	TS	PN	IL	RT	NP	QF	PL	PO	SG	LR	GI	TL	402
345	---	VD	TT	PD	TF	PH	LG	SI	Q	ANG	IG	SN	YI	---	GV	LS	SV	SP	---	---	---	---	---	378
403	NG	CH	DW	VG	T	PT	TW	PQ	---	QW	RC	SR	GS	NC	SS	GH	RY	---	VP	VM	NR	VT	VL	458
379	---	PS	H	SG	SQ	VD	LK	PI	NG	SS	IT	EA	TH	LA	PS	VY	PG	GE	VL	VV	FF	MS	KI	427
459	F	S	T	SR	---	LP	Q	LN	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	504
428	PG	AY	SL	P	CL	L	Q	EY	I	SH	LA	SE	QA	PT	VG	EA	LL	HY	VD	PT	GR	TL	GE	487
505	---	DN	I	V	M	P	NG	Y	R	FE	AW	NQ	F	Y	SL	A	P	VG	T	G	K	GR	---	538
488	AS	SG	P	Q	I	P	I	NG	V	F	V	SW	S	R	F	Y	Q	K	P	V	G	T	AS	523

RESULT 13

AAR57091
ID AAR57091 standard; protein: 530 AA.

XX	
AC	AAR57091;
XX	
DT	27-AUG-2003 (revised)
DT	25-MAR-2003 (revised)
DT	05-OCT-1994 (first en

Small round virus SRSV/KY/89 capsid protein.

pathogen; acute gastroenteritis; food poisoning; seafood contamination; diagnostic assay; human calcivirus; small round virus; SRSV; KY89; Norwalk virus; capsid protein.

Small round structured virus.

WO9405700-A2.

17-MAR-1994.

07-SEP-1993; 93WO-US008447.

07-SEP-1992; 92US-00941365.

(BAYU) BAYLOR COLLEGE MEDICINE.

Matson DO, Estes MK, Jiang X, Graham DY;

WPI; 1994-101125/12.

N-PSDB; AAQ56832.

DNA from Norwalk and related viruses - used for preparing prods. for use in diagnostic assays, detection and vaccines for Norwalk and related viruses.

Example 7; Fig 13a; 156pp; English.

CC The known sequence for Norwalk virus was used to obtain the sequence of
CC other Norwalk-related viruses such as SRSS/KY/89, an agent from a stool
CC outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide
CC cDNA sequence includes part of the polymerase region and the capsid
CC region of the genome; the deduced amino acid sequences are AAR57092 and
CC AAR57091, respectively. Expression of fragments and derivs. of Norwalk-
CC related viruses permits development of diagnostic assays to detect
CC antibodies, antigens, viral genetic material or antivirals. (Updated on
CC 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS
CC field.)

XX	Seq	Sequence 530 AA;
	Query Match	34.0%; Score 984.5; DB 2; Length 530;
	Beat Local Similarity	39.3%; Pred. No. 1.2e-81;
	Matches 230; Conservative	78; Mismatches 168; Indels 109; Gaps 16;
QY	1	MKMASNDAAAPSSDGAAG---LVPEIN-NVYMPLPVPAGASLATPVVGQQNIIDPWIRNNF 56
Dd	1	MMWASKDATSDVDGASASVLQVPEVNASDLAMDPAVAGSSTAVATAGVQNPIDPWIINNPF 60
QY	57	VQAPAGFTVSPRNSPGEIILLDELCPDIPNYPYLALHARWMYGAGGMVEQIVLAGNAFTA 116
Dd	61	VQAPQGFTTISPNNTPGDGVLFDSLGLPHLNPFLHLISQMYNGVGNMRVRIMLAGNAFTA 120
QY	117	GKIIFAAIPGCFPYENILSPSQITMCWHVIDVQLSPFLPMPDIWNPPHYNQGNDDPKL 176
Dd	121	GKIIVCIPGFGSQQLTIQAATFLPHVIADVTLDPIEVLDEVRNLPHNDRNQOTM 180
QY	177	RLVAMLYTPLRANNSGGDVFTVSCRVLTKPSDFEFLLVPPTVESKTQFALLPILKISE 236
Dd	181	RLVCMLYTPLSTGGGTGDSFWVAGRVMTCSPDNFLFLVPPTEVKTRPFTLPNLPSS 240
QY	237	MTNSRRFPVDVMYTARNENVQPQNGRVTLDGELLGTTPLLAVNICKEPGEVIAKNGD 296
Dd	241	LNSRSRAPLPISGMGISPDNVQSVQFONGRCTLDGRLVGTTTPVSLSHVAKIRG---TSNGT 297
QY	297	VRSYRMDEITNTDGTPTDPTEDTPGIGSPDFQG-----ILFGVASQRNKNEQNPAT 349
Dd	298	V-----INUTELDGTFFHFEG-PAPIGPDDLGGCDWHINMTQFHSSQTQYD----- 344
QY	350	RAHEAIINTGGDHLCPOISS-----SEIYLTSPNILRCNTNPQLPQSGLRGTLIRSD 402
Dd	345	-----VDTPPTSVPHLGSIQANGISGNYI---GVLSNVSP----- 378
QY	403	NGHCHDMVGTSPPTTWPO--QWRRCRSGNCCSSGHRYP----- 440
Dd	379	-----PSHPGSQVDLWKIKPNYGSSTIETHLAPSVSYPGFEVLVFPFMSKIPG 427
QY	441	-----VPVMNRVTWIVLSHKGSFSTSKLPLNLRWPLIRFINPDTRGVLFPEARLHKQ 495
Dd	428	PGGDSLCLLPQG---YISH-----LASEQAPTVG-EGPLLHVDPDTDNUGFEKRAYPD 478
QY	496	GFITVAHTG--DNPIVMPPNGYFRFEAWNQFYSLAPVGTGKGRR 538
Dd	479	GFLTCTVENGASSGPQQLPINGVFVFSWSRFRYQLKPEVGTASTAR 523
RESULT 14		
AAB49703	ID	AAB49703 standard; protein; 544 AA.
XX	AC	AAB49703;
XX	DT	04-APR-2001 (first entry)
XX	DE:	Small round structured virus protein SEQ ID 4.
KW		Small round structured virus; SRSV; food poisoning.
OS		Small round structured virus.
XX	FN	WO200079280-A1.
XX	PD	28-DEC-2000.
XX	PF	22-JUN-2000; 2000WO-JP004095.
XX	PR	22-JUN-1999; 99JP-00175928.
XX		(NTNA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA		(DENK-) DENKA SEIKEN KK.

```

PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI; 2001-080848/09.
DR N-PSDB; AAF29144.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
XX Claim 1; Page 47-49; 84pp; Japanese.
PS
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 544 AA;
SQ
Query Match 33.5%; Score 970; DB 4; Length 544;
Best Local Similarity 38.0%; Pred. No. 2.8e-80;
Matches 224; Conservative 91; Mismatches 176; Indels 98; Gaps 17;

QY 1 MKNASNDAAPSDGAAG---LVPEINN-EVMPLEPVAGASLATPVVQGNIIIDPWIRNMF 56
Db 1 MWMASKDAPSDGATGAGQLVPEVNTADPIRDPVAGSSTALATAGQVNLIDPWIRNMF 60
QY 57 VQAPAGEFTVSPNSPCEIILDLGLDLPNLYLAHARMYNGHAGMEVQIVLAGNAFTA 116
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLSQMYNGVGNMVRVVLVLAGNAFTA 120
QY 117 GKIIFAAIPGPFYENLSPSQTMCPIIIVRQLEPFLPMPDIIWNNFFHYNQNDPKL 176
Db 121 GKVIICCVFPGFQSRISIAQATLPHVIADVRLDPEVEPLEDVRNVLVH-NNDTQPTM 179
QY 177 RLVAAMYLTPLRAN--NSGDDVFTVSCRVLTKPSDPDEFTFLVPPTVESKTKQFALPLIKI 234
Db 180 RLLCMLYTLPLRTGGASGSDTSFVAGRVLTCPGDFNLFVLPPTVEQKTRPTVFNIP 239
QY 235 SEMTNSRFPVVDVMTARNENQVQPNQGRVTLTGELLGTTPLLAVINICKFKEVIAKN 294
Db 240 KYLSNRLIPNFIEMSLSPDQTVQFQNGRCTIDGQLGTTPTVSVSQLCKFRGR----- 294
QY 295 GDVRSYRMDMEITNTDGTPTDPTEDTGPIGSPDFQGLFGVASQRNKN---EQNP----- 347
Db 295 --ITSQRVLNLTLDGSPF--MAFAAPAPAGFPDLGSCDWHIEMSKIENSTQNNPIVTN 351
QY 348 -----ATRAHEALINTQGDHLCPOISSEIYLTSPNILRCTNPNQLPQSGLR 394
Db 352 SVKPSNQFVPHLSITILDENVSSGDVY-----GTLQWTS----- 388
QY 395 GTILIRSDNGHCHDMVGTSTPTTPWQWRRCRSGNCCSSGHRYPV--PVVMNRVTWIV 452
Db 389 -----PSUGGAN-----TNFKIPDYGSSLAESAQLAPVYPGFNEVIYVF 431
QY 453 LSHKSGFSTSTRK-----LPO-----LNLRW-----LIRFNPDTGRVLEARLHK 494
Db 432 MASIFPNQSGSNLVPCLLPQBYITHFISEQAP:QGEAALLHYVDPDTNRNLGEFKLVP 491
QY 495 QGFITVA--HTGDNPLVMPNPGYFRFEAWNOFYSLAPVGT--GKRRRV 540
Db 492 GGLYICVPSNSSSTGQQLDGLGVFVFNASVSRFYQLKPVGTAGPARGRL 540

RESULT 15
AAB49702
ID AAB49702 standard; protein; 546 AA.
XX
XX AAB49702;
XX
XX 04-APR-2001 (first entry)

```

```

XX Small round structured virus protein SEQ ID 3.
DE
XX Small round structured virus; SRSV; food poisoning.
KW
XX Small round structured virus.
OS
XX W0200079280-A1.
PN
XX 28-DEC-2000.
PD
XX 22-JUN-2000; 2000WO-JP004095.
PF
XX 22-JUN-1999; 99JP-00175928.
PR
XX (NTNA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
PA
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
PI
XX WPI; 2001-080848/09.
DR N-PSDB; AAF29143.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
XX Claim 1; Page 45-47; 84pp; Japanese.
PS
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 546 AA;
SQ
Query Match 33.4%; Score 967; DB 4; Length 546;
Best Local Similarity 38.5%; Pred. No. 5.3e-80;
Matches 225; Conservative 88; Mismatches 188; Indels 84; Gaps 17;

QY 1 MKNASNDAAPSDGAAG---LVPEINN-EVMPLEPVAGASLATPVVQGNIIIDPWIRNMF 56
Db 1 MWMASKDAPSDGATGAGQLVPEVNTADPIRDPVAGSSTALATAGQVNLIDPWIRNMF 60
QY 57 VQAPAGEFTVSPNSPCEIILDLGLDLPNLYLAHARMYNGHAGMEVQIVLAGNAFTA 116
Db 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLSQMYNGVGNMVRVVLVLAGNAFTA 120
QY 117 GKIIFAAIPGPFYENLSPSQTMCPIIIVRQLEPFLPMPDIIWNNFFHYNQNDPKL 176
Db 121 GKVIICCVFPGFQSRISIAQATLPHVIADVRLDPEVEPLEDVRNVLVH-NNDTQPTM 179
QY 177 RLVAAMYLTPLR--ANNNGDDVFTVSCRVLTKPSDPDEFTFLVPPTVESKTKQFALPLIKI 234
Db 180 RLLCMLYTLPLRTGGSGNSDSFVAGRVLTCTAATLPHVIADVRLDPEVEPLEDVRNVLVH-NNDTQPTM 239
QY 235 SEMTNSRFPVVDVMTARNENQVQPNQGRVTLTGELLGTTPLLAVINICKFKEVIAKN 294
Db 240 QTLNSRFPVVDVMTARNENQVQPNQGRVTLTGELLGTTPLLAVINICKFKEVIAKN 294
QY 295 GDVRSYRMDMEITNTDGTPTDPTEDTGPIGSPDFQGLFGVASQRNKN---EQNP----- 347
Db 295 --ITSQRVLNLTLDGSPF--MAFAAPAPAGFPDLGSCDWHIEMSKIENSTQNNPIVTN 351
QY 348 -----ATRAHEALINTQGDHLCPOISSEIYLTSPNILRCTNPNQLPQSGLR 394
Db 352 SVKPSNQFVPHLSITILDENVSSGDVY-----GTLQWTS----- 388
QY 395 GTILIRSDNGHCHDMVGTSTPTTPWQWRRCRSGNCCSSGHRYPV--PVVMNRVTWIV 452
Db 389 -----PSUGGAN-----TNFKIPDYGSSLAESAQLAPVYPGFNEVIYVF 431
QY 453 LSHKSGFSTSTRK-----LPO-----LNLRW-----LIRFNPDTGRVLEARLHK 494
Db 432 MASIFPNQSGSNLVPCLLPQBYITHFISEQAP:QGEAALLHYVDPDTNRNLGEFKLVP 491
QY 495 QGFITVA--HTGDNPLVMPNPGYFRFEAWNOFYSLAPVGT--GKRRRV 540
Db 492 GGLYICVPSNSSSTGQQLDGLGVFVFNASVSRFYQLKPVGTAGPARGRL 540

RESULT 15
AAB49702
ID AAB49702 standard; protein; 546 AA.
XX
XX AAB49702;
XX
XX 04-APR-2001 (first entry)

```

Db 383 TIWISQSTPPGTDINLWEIPDYGSSLSQAANLAP-PVPPPGFGEALVYFVSAPPGPNN 441
Qy 465 K-----LPQ-----LNLRW-----LIRFINPDTGRVLFPEARLHKOGFITVA-- 501
Db 442 RSAPNDVPCLLPQEIYTHFVSEQAPTMGDALLHYVDPDTNRNLGFEKLYPGGYLTCVEN 501
Qy 502 HTCDNFIVMPNPGYFRFEAWVQFYSLAPVGTGK-----GRRV 540
Db 502 GVGAGPQLPLNGVFLFVSWSRFYQLKPVGTASTARSLGVRI 546

Search completed: June 1, 2004, 13:46:03
Job time : 48.4752 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 12.8896 Seconds
(without alignments)
2166.837 Million cell updates/sec

Title: US-09-926-799-11
Perfect score: 2894
Sequence: 1 MKNASNDAPSSDGAAGLVP.....VNFYSLAPVGTGKRRRVQ 541

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/prodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCTUS COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1008.5	34.8	530	4	US-08-486-049-3
2	267.5	9.2	622	4	US-09-590-020-6
3	265.5	9.2	623	4	US-09-590-020-2
4	265.5	9.2	623	4	US-09-590-020-4
5	264	9.1	668	4	US-09-617-594A-4
6	262.5	9.1	626	4	US-09-590-020-7
7	250.5	8.7	669	4	US-09-617-594A-2
8	114	3.9	630	4	US-09-396-149-5
9	114	3.9	2206	1	US-07-852-260-2
10	114	3.9	2206	2	US-08-461-503-2
11	114	3.9	2206	3	US-08-465-250-2
12	109.5	3.8	1101	3	US-08-986-485-2
13	104.5	3.6	2318	3	US-09-091-219-24
14	104.5	3.6	2318	4	US-09-660-541-24
15	102.5	3.5	596	4	US-09-489-039A-9905
16	101	3.5	1013	3	US-09-415-522-8
17	96	3.3	803	4	US-09-543-681A-4312
18	96	3.3	1082	1	US-08-106-493A-2
19	96	3.3	1082	2	US-08-429-264-2
20	96	3.3	1139	1	US-08-832-883-2
21	96	3.3	1139	2	US-08-832-877-2
22	95	3.3	439	4	US-09-543-681A-5727
23	95	3.3	1050	4	US-09-428-711A-16
24	93.5	3.2	1043	3	US-08-928-361B-30
25	93.5	3.2	1721	3	US-08-700-651-5
26	93.5	3.2	1721	3	US-08-928-361B-6
27	93.5	3.2	1721	4	US-09-588-995A-6

28	93	3.2	1162	3	US-08-803-346-1	Sequence 1, Appli
29	92.5	3.2	1560	4	US-09-264-512B-2	Sequence 2, Appli
30	92	3.2	589	4	US-09-543-681A-6155	Sequence 6155, Ap
31	91	3.1	895	3	US-08-827-962-19	Sequence 19, Appl
32	91	3.1	1162	3	US-08-827-962-15	Sequence 15, Appl
33	90.5	3.1	248	4	US-09-540-236-3073	Sequence 3073, Ap
34	90.5	3.1	1091	3	US-08-986-485-5	Sequence 5, Appli
35	90.5	3.1	1911	4	US-09-854-856-64	Sequence 64, Appli
36	90.5	3.1	1939	4	US-09-854-856-48	Sequence 48, Appl
37	90.5	3.1	1971	4	US-09-854-856-32	Sequence 32, Appl
38	90.5	3.1	1999	4	US-09-854-856-16	Sequence 16, Appl
39	90.5	3.1	2004	4	US-09-854-856-58	Sequence 58, Appl
40	90.5	3.1	2032	4	US-09-854-856-42	Sequence 42, Appl
41	90.5	3.1	2048	4	US-09-854-856-62	Sequence 62, Appl
42	90.5	3.1	2064	4	US-09-854-856-26	Sequence 26, Appl
43	90.5	3.1	2076	4	US-09-854-856-46	Sequence 46, Appl
44	90.5	3.1	2092	4	US-09-854-856-10	Sequence 10, Appl
45	90.5	3.1	2108	4	US-09-854-856-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1
US-08-486-049-3
; Sequence 3, Application US/08486049
; Patent No. 6572862
; GENERAL INFORMATION:
; APPLICANT: Bates, Mary K
; APPLICANT: Jiang, Xi
; APPLICANT: Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Ave., N.W.
; CITY: Washington, D.C.
; STATE:
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,049
; FILING DATE: June 7, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311.023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-486-049-3

Query Match 34.8%; Score 1008.5; DB 4; Length 530;
Best Local Similarity 40.9%; Pred. No. 4.2e-96;
Matches 235; Conservative 72; Mismatches 180; Indels 87; Gaps 14;
QY 1 MKNASNDAPSSDGAAG---LVPEIN-NEVMPLEPVAGSLATPVVQGNIIDPWIRNPF 56
DB 1 MMWASKDATSSVDGASGAGQLVPEYNASDPLAMDPAVGSSTAVATAGVNPIDPWIRNPF 60

7

;; TITLE OF INVENTION: CDNA
;; NUMBER OF SEQUENCES: 9
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Cooper & Dunham LLP
;; STREET: 1185 Avenue of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: U.S.A.
;; ZIP: 10036
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release 1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/465,250
;; FILING DATE: 6-JUN-1995
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: White, John P.
;; REGISTRATION NUMBER: 28,678
;; REFERENCE/DOCKET NUMBER: 36607-E-PCT-US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 278-0400
;; TELEFAX: (212) 391-0525
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2206 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-465-250-2

Query Match 3.8%; Score 114; DB 3; Length 2206;
Best Local Similarity 18.1%; Pred. No. 0.099;
Matches 136; Conservative 86; Mismatches 237; Indels 294; Gaps 34;

Qy 26 VMLEPV---AGASLATPVV-----GQNIIDP-----WIENNF 56
Db 299 ILPLSLDFAQSSVEIPITVTIAPMCSFNGRLNVTAPKQGLPVLNTPGNSOYLTSN 358
Qy 57 VQAPAG--EFTVSPR--NSPGRI-----75
Db 359 HQSPCALPEFDVTPIDIPGVEKRMELAEIDTMIPLNLSKENTMDYRVTLSDSADL 418
Qy 76 ---LLDLELGLDLPYLAAH-----LARMYNGHAGMEVQIVLAGNAFTAKIIPAAIPPG 127
Db 419 SQPILCLSLSPAFDPRLSHTMLGVLNYTHWAGSLKFTFLFCGMMATGKILVAYAPPG 478
Qy 128 FPYENLSPQITMCPHVIIIVRQLEPFLPMPDIWNFFHYNQNDP-----K 175
Db 479 -AQPTSRKEAMLGTHVIMDLGLOSSCTMVPF--WISNVYRQTQDSFTGGVISMFPYQ 535
Qy 176 LRLVAMLYTPLRANSGDDVFTVSCRVLTKSPDPFEFTL-----VPPTVESKTK 225
Db 536 TRIVVPLSTPKSMMLG---FVSACN-----DFSVELLADTTHISQALPQIGEDLTS 585
Qy 226 QFA-----LPILKISEMTNSR--FP-----VPDVMYT-----251
Db 586 EVAQCALTLSPKQDLSLPDTKASGPAHSKEVPALTAVETGATNPLAPSDTVQTRHVVR 645
Qy 252 -ARNEN-----QVQPQNGRVTLDGELLGTPLLAVNICKFKGEV-IAGKGDV 297
Db 646 RSRSESTIESFPARGACVAIIEVDNEQPTTAAQ-----KLFAMWRITVKTQVLRKLEF 700
Qy 298 RSY-RMDME-----ITNTD-----GTPIDPTED-----319
Db 701 FYISRFDEFTFVVTANTFNANNGHALNQVYIMVPPGAPTPKSDDDYTWTQSSNFSIF 760
Qy 320 -----TPGPIGSPDQGIH-----FGVASQ 339
Db 761 YTYGAAPARISVP-YVGLANAYSHFYVDGPAKVELKTDANDQIGDSLISAMTVDDFGVLAV 819

Qy 340 RNKNEONPATRAHEAIINTGGDHL---CPQISSSEIYLTSPNILRCTNPQLPQSGLRGT 396
Db 820 RVVDHNPTKVTSKVRIYMKPKHVRVWCPRPRAVPYI--GPGVDYRNNDLPSEKG---- 874
Qy 397 ILIRSDNGHCHDMVGTSPPTTTPWPOQWRCRSGSNCCS-----SGHRYVPVVMNRVTW 450
Db 875 -LITYGFGHONKAVYTA-----GYKICNYHLATKEDLQNAVSIMWNRDLL 918
Qy 451 IVLSHKSFGSTSTRKLPQLNLRWPLIRFINPDTGRVLFEARLHKQGFITVAHTGDNPIVM 510
Db 919 VVESKAQGTDSIAR-----CNCNAGVYICESR-----RKYPVSVFGVPTFOYM 961
Qy 511 PPNGYR--RFEAWNOFYSLAPVGTGKRRRVQ 541
Db 962 EANDYYPARYQSHMLIGHGFASPDGCGILRCQ 994

RESULT 12
US-08-986-485-2
; Sequence 2, Application US/08986485
; Patent No. 6046030
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUNEH, ALEMSEGED
; TITLE OF INVENTION: A HUMAN LIG-1 HOMOLOG (HLIG-1)
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/986,485
; FILING DATE: 08-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/059,448
; FILING DATE: 22-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1101 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-986-485-2

Query Match 3.8%; Score 109.5; DB 3; Length 1101;
Best Local Similarity 21.1%; Pred. No. 0.09;
Matches 120; Conservative 55; Mismatches 193; Indels 201; Gaps 33;

Qy 51 WIRNFWQAPAG-EFTVSPRNSPGETLLDLELGPDLNRY--LAHLARMYNGHAGMEVQI 107
Db 538 WKKDNEVLNTADMENFVHQAQGEVW-----EYTTILHLRQVTFGHEG--RYQC 585
Qy 108 VLA---GNAFTAGKIIFAAIPGFFPYENLSPSQITMCPHVI-----IDVRQLEPFL--PM 158

```
Db 586 VITNFGSTYSHKARLTVMNLSF-----TKTPHDITRTTTTARLECAATGPHN 635
QY 159 PDI-MNNFFHYNOGND-PKLRVLVAMLYTLRANNSGDDVFTV-----SC----- 200
Db 636 PQIAWQK-----DGGTDFPAARERRHVMPP-----DDVFFITDVKIDDAGVYSCAONS 685
QY 201 -----RVLTQPS-----PQFETFLVPPTVESKTKQFALPKIKSEMNTSRFPVVD 247
Db 686 AGSISANATLTVLETPSLVPLEDRVVSGETVALQCKATGNPPPRITFWFKGDR-PLSLT 744
QY 248 VMYTARNENOVQPON-----GRVTLG-GEILGT-----TPLLAVNICKEGFEVIKN 294
Db 745 ERHHLTPDNLVQNVQVAEDAGRYCEMSNTLGTERRAHSQSLVLPAAACRKGDTTV--- 801
QY 295 G-----DVRSYRMDMEITNDGTPIDPTEDTGPIGSPDFQGI 332
Db 802 GIPTIAVSSIVLTSVWCIIYQTRKKEEYSVNTDVTVP--DVPSYLS--QQT 856
QY 333 LFGVASQRNKNQNPATRAHEAIINTGDLCPQISSSEIYLTSPNI--LRCNTPQ----- 386
Db 857 LSDROETVVRTEGGPOANGH--IESNG--VCPRDAS--HFPEDTHSVACRQPKLCAG 908
QY 387 -----PLPQSLRGTILIRSDNGHCHDMV-----CT 412
Db 909 SAYHKEPWKAMEKAEGTGPHEKHGGRVYVCSN---CNTEVDCYSRQAFHPOPVSRSDA 965
QY 413 SPTTPTWPO-----QWRRCSRGSNCCSSGHRYPVPVWNRVTWIVLSHKSQFS 460
Db 966 QPSAPNGPFGGSDQEHSPHQCSTAAAGSCPECQSLP-----SNHDMRL 1012
QY 461 TSTRKLPLQNL-----RWPLIRFINPDT 483
Db 1013 TAVKKKPMASLDGKGDSSWTLARLYHPS 1041

RESULT 13
US-09-091-219-24
; Sequence 24, Application US/09091219
; Patent No. 6171592
; GENERAL INFORMATION:
; APPLICANT: STUDDERT, Michael J.
; APPLICANT: CRABB, Brendan S.
; APPLICANT: FENG, Li
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
; CURRENT APPLICATION NUMBER: US/09/091,219
; PRIOR FILING DATE: 1998-10-05
; EARLIER FILING DATE: 1996-12-18
; EARLIER FILING DATE: 1995-12-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 2318
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
US-09-091-219-24
Query Match 3.6%; Score 104.5; DB 3; Length 2318;
Best Local Similarity 22.6%; Pred. No. 1.1;
Matches 33; Conservative 21; Mismatches 67; Indels 25; Gaps 2;

QY 79 LEIGPDLNPLYLAHARMYNGHAGGMEVQIVLAGNAFTAGKIIFAALPGGFPYENLSPSQI 138
Db 367 LELPTDHKGVYGLTDSYAYMRGNDVETAVGNQFNGCGLLVAMVPELYSIQKRELYQL 426
QY 139 TWCPHVIVDROLEPFLPMPDIWNNFFHYNOGNDPKLRVLVAMLYTLRANNSG----- 192
Db 427 TLFPHQFINPRTNMTAHITVPFGVGNRYDQYKHKP-WTLVVMVAVPLTVNTGAPQIKV 485
QY 193 -----DDVFTVSC 200

US-09-091-219-24
Query Match 3.6%; Score 104.5; DB 3; Length 2318;
Best Local Similarity 22.6%; Pred. No. 1.1;
Matches 33; Conservative 21; Mismatches 67; Indels 25; Gaps 2;

QY 79 LEIGPDLNPLYLAHARMYNGHAGGMEVQIVLAGNAFTAGKIIFAALPGGFPYENLSPSQI 138
Db 367 LELPTDHKGVYGLTDSYAYMRGNDVETAVGNQFNGCGLLVAMVPELYSIQKRELYQL 426
QY 139 TWCPHVIVDROLEPFLPMPDIWNNFFHYNOGNDPKLRVLVAMLYTLRANNSG----- 192
Db 427 TLFPHQFINPRTNMTAHITVPFGVGNRYDQYKHKP-WTLVVMVAVPLTVNTGAPQIKV 485
QY 193 -----DDVFTVSC 200
```

```
Db 486 YANIAPTNNHVAGBPPSKGPIPVAC 511

RESULT 14
US-09-660-541-24
; Sequence 24, Application US/09660541
; Patent No. 6531136
; GENERAL INFORMATION:
; APPLICANT: STUDDERT, Michael J.
; APPLICANT: CRABB, Brendan S.
; APPLICANT: FENG, Li
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
; FILE REFERENCE: 040268/0151
; CURRENT APPLICATION NUMBER: US/09/660,541
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/091,219
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: AU PN7201
; PRIOR FILING DATE: EARLIER FILING DATE: 1995-12-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 2318
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
US-09-660-541-24
Query Match 3.6%; Score 104.5; DB 4; Length 2318;
Best Local Similarity 22.6%; Pred. No. 1.1;
Matches 33; Conservative 21; Mismatches 67; Indels 25; Gaps 2;

QY 79 LEIGPDLNPLYLAHARMYNGHAGGMEVQIVLAGNAFTAGKIIFAALPGGFPYENLSPSQI 138
Db 367 LELPTDHKGVYGLTDSYAYMRGNDVETAVGNQFNGCGLLVAMVPELYSIQKRELYQL 426
QY 139 TWCPHVIVDROLEPFLPMPDIWNNFFHYNOGNDPKLRVLVAMLYTLRANNSG----- 192
Db 427 TLFPHQFINPRTNMTAHITVPFGVGNRYDQYKHKP-WTLVVMVAVPLTVNTGAPQIKV 485
QY 193 -----DDVFTVSC 200

Db 486 YANIAPTNNHVAGBPPSKGPIPVAC 511

RESULT 15
US-09-489-039A-9905
; Sequence 9905, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9905
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9905
Query Match 3.5%; Score 102.5; DB 4; Length 596;
Best Local Similarity 21.4%; Pred. No. 0.17;
Matches 112; Conservative 52; Mismatches 221; Indels 139; Gaps 24;

QY 99 HAGGMEVQIVLAGNAFTAGKIIFAALPGGFPYENLSPSQITMCP-----HVIDVRQLRP 153
Db 52 HAGCARVECVIADTCI-AGMAESAACEKFSNQNGVT--ITVPCWCYGSSETIDMDPLRP 109
```

Search completed: June 1, 2004, 13:58:11
Job time : 14.8896 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 34.5841 Seconds

(without alignments)
4368.312 Million cell updates/sec

Title: US-09-926-799-11

Perfect score: 2894

Sequence: 1 MKMASNDAPSDGAGLVP.....VNQFYS LAPVTGKGRRRVQ 541

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCT05_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1008.5	34.8	530	14	US-10-314-739-3
2	264	9.1	668	14	US-10-209-507-4
3	262.5	9.1	547	12	US-10-670-695-4
4	262.5	9.1	671	12	US-10-670-695-2
5	250.5	8.7	669	14	US-10-209-507-2
6	118.5	4.1	1147	15	US-10-327-481A-38
7	114	3.9	630	14	US-10-372-686-5
8	114	3.9	630	14	US-10-371-558-5
9	114	3.9	630	14	US-10-375-553-5
10	114	3.9	630	16	US-10-372-553-5
11	108.5	3.7	861	12	US-10-282-122A-66008
12	108	3.7	3696	16	US-10-229-148B-4
13	107	3.7	607	12	US-10-258-951-74
14	106.5	3.7	861	12	US-10-296-115-1211
15	106.5	3.7	861	12	US-10-282-122A-65330

16	106.5	3.7	1077	15	US-10-093-463-136	Sequence 136, Appl
17	103	3.6	578	9	US-09-924-358-39	Sequence 39, Appl
18	103	3.6	578	15	US-10-410-764-39	Sequence 39, Appl
19	102.5	3.5	656	14	US-10-128-714-3343	Sequence 3343, Ap
20	102.5	3.5	1049	12	US-10-211-462-6	Sequence 6, Appl
21	102.5	3.5	1049	12	US-10-170-385-379	Sequence 379, Appl
22	102.5	3.5	1049	15	US-10-236-031B-52	Sequence 52, Appl
23	102.5	3.5	1049	15	US-10-295-027-4	Sequence 4, Appl
24	102.5	3.5	1049	15	US-10-295-027-726	Sequence 726, Ap
25	102.5	3.5	1049	15	US-10-295-027-1166	Sequence 1166, Ap
26	102.5	3.5	1049	15	US-10-301-330-2	Sequence 2, Appl
27	102.5	3.5	1049	15	US-10-188-832-2	Sequence 2, Appl
28	102.5	3.5	1063	12	US-10-425-114-39245	Sequence 39245, A
29	101.5	3.5	639	12	US-10-424-599-144381	Sequence 144381, A
30	100.5	3.5	723	12	US-10-425-114-38957	Sequence 38957, A
31	100.5	3.5	837	9	US-09-738-626-6913	Sequence 6913, Ap
32	99.5	3.4	745	14	US-10-128-714-8343	Sequence 8343, Ap
33	98.5	3.4	481	15	US-10-389-566-826	Sequence 826, Appl
34	98.5	3.4	1936	12	US-10-152-886-1	Sequence 1, Appl
35	98.5	3.4	2781	12	US-10-263-929-122	Sequence 122, Appl
36	98	3.4	564	14	US-10-156-761-8951	Sequence 8951, Ap
37	98	3.4	3931	15	US-10-120-801-18	Sequence 18, Appl
38	97.5	3.4	4952	15	US-10-051-874-56	Sequence 56, Appl
39	97.5	3.4	5008	15	US-10-051-874-166	Sequence 166, Appl
40	97.5	3.4	5159	15	US-10-085-198-112	Sequence 112, Appl
41	97	3.4	865	12	US-10-282-122A-77101	Sequence 77101, A
42	97	3.4	2515	12	US-10-042-865-53	Sequence 53, Appl
43	97	3.4	2515	12	US-10-029-020-53	Sequence 53, Appl
44	97	3.4	6310	12	US-10-282-122A-67793	Sequence 67793, A
45	96.5	3.3	262	14	US-10-059-909-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-10-314-739-3
; Sequence 3, Application US/10314739
; Publication No. US20030129588A1
; GENERAL INFORMATION:
; APPLICANT: Bates, Mary K
; Jiang, Xi
; Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; Characterize No. US20030129588A1walk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pubbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Ave., N.W.
; CITY: Washington, D.C.
; STATE: <Unknown>
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/314,739
; FILING DATE: 09-Dec-2002
; CLASSIFICATION: <Unknown>
; PRIORITY INFORMATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,049
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311.023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX: <Unknown>

```

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 530 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-314-739-3

Query Match          34.8%; Score 1008.5; DB 14; Length 530;
Best Local Similarity 40.9%; Pred. No. 6.2e-87;
Matches 235; Conservative 72; Mismatches 180; Indels 87; Gaps 14;

QY 1 MKVASNDAAAPSDGAG---LVEIN-NEVMLEPVAGASLATPVVQGNIIIDPIRNRF 56
DB 1 MMASKDATSVDGASGAGLVPEVNASDPLAMDVPAGSLTAVATAGVNPIDPWIIINF 60
QY 57 VQAPAGEFTVSPRNSGEILLDLELGDPLNPYLALHARMYNGHAGMEVQIVLAGNAFTA 116
DB 61 VQAPQGEFTISPNTPGDVLFDLSLGPLNPLFLLHLSQYINGWGNMVRIMLAGNAFTA 120
QY 117 GKLIIFAAIPGFPYENLSPSQITMCPHVIDVROLEPFLPMPDIWNFFHYNQGNPKL 176
DB 121 GKLIIVCIPPGFSGHNLITAAQLFPHVIADVRLDPIEVPLEDVRLVLFHNNDRNQTM 180
QY 177 RLVMALYTPLRANSGDDVFTVSCRVLTKPSDPFETFLVPPTVESKTKQFALPILKISE 236
DB 181 RLVCMLYTPLRITGGTGDSFVAVAGRYMTCPSDPDFLFLVPPTVEQKTRPFTLPLSLSS 240
QY 237 MTNSRPVPVDMVTARNENQVQPONGRVTLTDLGELLGTPPLAVNICFKFGEVIAKNGD 296
DB 241 LNSRAPLPISMGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRG---TSNGT 297
QY 297 VRSYRDMETINTDGTPTDPTEDTPGPIGSPDFQG-----LFGVASQRNKNQNPAT 349
DB 298 V-----INLTEDGTFPHFEG-PAPIGFPDLGGCDWHINMTQFGHSSQTQYD----- 344
QY 350 RAHEAINTGDHLCPQISSEIYVILSPNLRCTNPQLPQSGLRGTILIRSDNGCHDM 409
DB 345 -----VDITDPTDFVPLHGS-----IQANGISGNY 369
QY 410 VG-----TSPTTPTWPO--QWRRCRSGNSCCSSGHRYP--VPVVMNRVTWIVLSHSGFS 460
DB 370 VGVLSWISPSHSGSGVDLWKIPNYGSSITEATHLAPSVYPGFGVLVFEFMSKMPGP 429
QY 461 TSTRK--LPQLNL-----RWPLIRFNPDTGRVLFEARLHKQGFITVAHTG-- 504
DB 430 AYNLPCLLPQEIYISHLASEQAPTVEAALLHYVDPDTGRNLGEFKAYPDGFLTCVPNGAS 489
QY 505 DNPIWPPNGVFRFANVQFSLAPVGTGKGR 538
DB 490 SGPOQLPINGVFVFSVWSREYQLKPVGTASSAR 523

RESULT 2
US-10-209-507-4
; Sequence 4, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V
; FILE REFERENCE: 454313-3151-2
; CURRENT FILING DATE: 2002-11-12
; PRIOR FILING DATE: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR FILING DATE: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR FILING DATE: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PR
; ORGANISM: Feline calicivirus
; US-10-209-507-4

Query Match          9.1%; Score 264; DB 14; Length 668;
Best Local Similarity 38.0%; Pred. No. 8.7e-16;
Matches 60; Conservative 25; Mismatches 59; Indels 14; Gaps 5;

QY 67 SPRNSPGEILLDLELGDPLNPYLALHARMYNGHAGMEVQIVLAGNAFTAGKIIFAAIPP 126
DB 179 STSETQCKILFKQSLGPLNPLYLTHLAKLYVAVWSGSIEVRSISGSGVFGKLAIVVPP 238
QY 127 GF-PYENLSPSQITMCPHVIDVROLEPFLPMPDIWNFFHYNQGNPKLRLVAMLYTP 185
DB 239 GIDPVQSTMLQY---PHVLFDARQVFEVFTIPDLRNSLYHLMSDDTD-TTSLVIMLYND 294
QY 186 L-----RANNSGDDVFTVSCRVLTKPSDPFETFLVPP 218
DB 295 LINPYANDNSSGCIVIVE---TKPGDPDFKFLKPP 328

RESULT 3
US-10-670-695-4
; Sequence 4, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 547
; TYPE: PR
; ORGANISM: Feline calicivirus
; US-10-670-695-4

Query Match          9.1%; Score 262.5; DB 12; Length 547;
Best Local Similarity 25.3%; Pred. No. 8.9e-16;
Matches 113; Conservative 59; Mismatches 167; Indels 107; Gaps 20;

QY 30 EPVAGASLATPVVQGNIIIDPIRNFPVQAPAGEFTV---SPRNSPGEILLDLELGPOLN 86
DB 21 EPSAQNSTAADATKGSVDSEW-----EAFPSFHTSVNMWSTSETQCKILFKQSLGPLN 74
QY 87 PYLAHLARMYNGHAGMEVQIVLAGNAFTAGKIIFAAIPPGF-PYENLSPSQITMCPHVI 145
DB 75 PYLEHLAKLYVAVWSGSIEVRSISGSGVFGKLAIVVPPGVDPVQSTMLQY---PHVL 131
QY 146 IDVRQLEPFLPMPDIWNFFHYNQGNPKLRLVAMLYTPLRANSGDDDVFTVSC--RVL 203
DB 132 FDAQVEPVI FCLPDLRLSTLYHLMSDDTD-TTSLVIMVYNDL-INPYANDANSGCCIVTVE 189
QY 204 TKPSDPFETFLVPPTVESKTKQFALPILKISEMNSRFP---VP-----VDVMYTARN 254
DB 190 TKPGDPFKFLKPPG-----SMLTHGSIPLPKTSSLWIGNRYWSDI 234
QY 255 ENQVQVQ-----QNGRVTLDGELLG-TTPL-----LAVNICFKFGEVIAKNGDVRSYR- 302
DB 235 TDFVIRPFVQANRHFDNQETAGWSTPRFRPISVTITEQNGAKLG-IGVATDIYVPGIP 293
QY 303 -----DMEITNDGTPIDPTD-----DTPGPI-GSPDPFGGILFGVASQRNKN 343
```

Db 294 DCGPDTTIPGELIPAGDYAITNGTNDITATGYDTADIIKNNTPRGMVYCGSLQRAWG 353
Qy 344 EQNPATRAHEALINTCGD-----HLCPOISSSEIYL-----TSPNLL----- 380
Db 354 DKKISNTAFITATLDGNNKINPCNTIDQSKI VVFDQNHVHGKKAQTSDDTLALLGYTG 413
Qy 381 -----RCTNPQPLPQSGLRG 395
Db 414 ICEQAIGSDRDRVVRISTLTPETGARG 439

RESULT 4
US-10-670-695-2
; Sequence 2, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 671
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-670-695-2

Query Match 9.1%; Score 262.5; DB 12; Length 671;
Best Local Similarity 25.3%; Pred. No. 1.2e-15;
Matches 113; Conservative 59; Mismatches 167; Indels 107; Gaps 20;

Qy 30 EPVAGSLATPVVGQNIIDPIRNNFVOAPAGEFTV---SPRNSPGEILLDLGLGPDLN 86
Db 145 EFSAQMSTAADATGATKSVDSW-----EAFSFTSVNWSSETQGKILFKQSLGPLLN 198
Qy 87 PYLAHLARMYNGHAGMEVQIVLAGNAFTAGKIIFAAIPPGF-PYENLSPSQITMCPHVI 145
Db 199 PYLEHLAKLYVAWSGSIERFSISGSGVFGKLAALVVPVPGDPVQSTSMLOY---PHVL 255
Qy 146 IDVRQLEPPELLPMPDIWNFFHYNQGNPKLRLVAMLYTTPLRANSGDDVFTVSC--RVL 203
Db 256 FQARQVEPVIFCLPDLRSTLYHLMSDTP-TTSLVIMVYNDL-INPYANDANSNGCIVTVE 313
Qy 204 TKPSDPDEFTEFLVPPTVESKTKQFALPILKISEMTNSREP---VP-----VDVMYTARN 254
Db 314 TKPGDPFKFHLKPPG-----SMLTHGSIPSLDLPKTSLSLWNGRYSWI 358
Qy 255 ENQVQVQ---QNGRVTLDGELLG-TTPL---LAVNICFKGEVIKAGNDVRSYRM---- 302
Db 359 TDFVIRPFVQANRHFDFNQETAGSTPRFRISVITTEQNGAKLG-IGVATDIYVPGIP 417
Qy 303 -----DMEITNDGTPIDPTE--DTPGPI-GSPDPFGILFGVAGSQNNK 343
Db 418 DCGPDTTIPGELIPAGDYAITNGTNDITATGYDTADIIKNNTPRGMVYCGSLQRAWG 477
Qy 344 EQNPATRAHEALINTCGD-----HLCPOISSSEIYL-----TSPNLL----- 380
Db 478 DKKISNTAFITATLDGNNKINPCNTIDQSKI VVFDQNHVHGKKAQTSDDTLALLGYTG 537
Qy 381 -----RCTNPQPLPQSGLRG 395
Db 538 ICEQAIGSDRDRVVRISTLTPETGARG 563

RESULT 5
US-10-209-507-2
; Sequence 2, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-2

Query Match 8.7%; Score 250.5; DB 14; Length 669;
Best Local Similarity 25.6%; Pred. No. 1.7e-14;
Matches 93; Conservative 58; Mismatches 141; Indels 71; Gaps 15;

Qy 67 SPRNSPGEILLDLGLGPDLPNYPYLAHLARMYNGHAGMEVQIVLAGNAFTAGKIIFAAIPPP 126
Db 180 STSETQGKILFKQSLGPLNLYLEHLSKLYVAWSGSDVRSISGSGVFGKLAALVWPP 239
Qy 127 GF-PYENLSPSQITMCPHVIIDVRQLEPPELLPMPDIWNFFHYNQGNPKLRLVAMLYTP 185
Db 240 GVDVPQSTSMLOY---PHVLFQARQVEPVIFSIPLRSTLYHLMSDTP-TTSLVIMVYND 295
Qy 186 L-----RANSGDDVFTVSCRVLTKPSPDFEFTPLVPPTVESKTKQFALPILKISEMTNS 240
Db 296 LINPYANDANSNGCIVTVE---TKPGDPFKFHLKPPG-----SMLTHG 336
Qy 241 RPP---VP-----VDVMYTARNQVQVQ---QNGRVTLDGELLG-TTPL---LAVNI 283
Db 337 SIPSDLPILKSSLSLWNGRYSWSDITDFVIRPFVQANRHFDFNQETAGWSTPRFRPITITI 396
Qy 284 CKFKGEVIKAGN-----GDVRSYRMDMEITNDGTPIDPTE--DTPG 322
Db 397 SESNGSKLGTGVATDIYVPGIPDGWPDITIGELTPAGDYSITNGSGNDIATANAYDSAD 456
Qy 323 PT-GSPDPFGILFGVAGSQNNKNEQNPATRAHEALINTGDLHLCPOISSSEIYLTSPNLR 381
Db 457 VITNTNFRGMVYCGALQRAWGDKKISSTAFITTAIKSGNTLKP---SNTIDMTKIAVVQ 513
Qy 382 CTN 384
Db 514 DTH 516

RESULT 6
US-10-327-481A-38
; Sequence 38, Application US/10327481A
; Publication No. US20040001864A1
; GENERAL INFORMATION:
; APPLICANT: King, Andrew M.Q.
; APPLICANT: Burman, Alison J.
; APPLICANT: Audonnet, Jean-Christophe F.
; APPLICANT: Lombard, Michel F.A.
; TITLE OF INVENTION: Vaccine against Foot-and-Mouth Disease
; FILE REFERENCE: 454313-3178
; CURRENT APPLICATION NUMBER: US/10/327,481A
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: PCT/FR01/02042

Wed Jun 2 09:13:25 2004

```

; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: FR 0008437
; PRIOR FILING DATE: 2000-06-29
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 1147
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
US-10-327-481A-38

Query Match
Best Local Similarity 4.1%; Score 118.5; DB 15; Length 1147;
Matches 77; Conservative 55; Mismatches 156; Indels 137; Gaps 13;

QY 32 VAG---ASLAPVVGQNIIDPWIRNFFVQAPAGETVSPRNSPGBEILDLELGPDLNPPY 88
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 129 VAGPNTSGLETRVVQAEFFKFLPDWTTDKPFGYLT-----KLELPTDHHGV 176

QY 89 LAHLARMYNGHAGMEVQIVLAGNAFTAGKIIFAAIPGPPYENLSPSQITMCPHVLDV 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 177 FGLVDSYAYMRNGWDVEVSAVGNQFNGGCLLVAMVPEWKAFTREKYQLTLFPHQFISP 236

QY 149 RQLEPILLPMDIWNFFEHVNGQNDPKLRVAMLYTPLRANN----- 190
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 237 RTNMTAHTITVYLGVRNRYDQYKKHKP-WTLVVMVLSPLTVSNTAAPQIKVYANIAPTYVH 295

QY 191 -----SGDDVFTVSC-----RVLTGPSDF--BFTFLVPPTVES 222
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 296 VAGELPSKEGIFPVACADGGLVTTDPKTADPVYGVKNVPPKTNYPGRFTNLLDAEAC 355

QY 223 KT-----KQFALPILKISEMNSRPPVVDVWYARNEN-----QVVPQNGRVT 267
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 356 PTELRFDGKPYV-----VTRADDTRLAKFDVSLAAKMSNTYLSGIAQYTYQYSGTIN 410

QY 268 LDGELGTTPLLVANICKRKEGIKNGDVSRYMDMEITNDGTPTDPTDPTDPTDPTDPT 327
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 411 LHMFTGSTDKARYWVAY-----IPPGVEITP-----P 438

QY 328 DFOGILFGVASQRNKNQNPATRAH-----EALINTGGDHLCPQISSSEIYLTSPNLR 381
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 439 D-----TPEEAHCIIHAEDWTGLNSKFTFSIPYVSAADYAYTASDTAE 481

QY 382 CTNPQ 386
   ||| |||
Db 482 TTNVQ 486

RESULT 7
US-10-372-686-5
; Sequence 5, Application US/10372686
; Publication No. US20030159185A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Replication Protein A and Use
; FILE REFERENCE: 0899D
; CURRENT APPLICATION NUMBER: US/10/372,686
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/396,149
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/123,896
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/100,690
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-372-686-5

Query Match
Best Local Similarity 3.9%; Score 114; DB 14; Length 630;
Matches 67; Conservative 42; Mismatches 99; Indels 112; Gaps 16;

QY 123 AIPPG---FPYENLSPSQIT--MCPHVLDVRLQLEPFLFLLPMDIWNPF--FHYNQNDP-K 175
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8 SVTPGAVAFVLENASPDAAATGVPVEIVLQVVDLKP-----IGTRFTFLASDGDKIK 60

QY 176 LRLVAMLYTPLRANNSGD---DVFTVSCR-----VLTGPSDFEF----- 212
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 TMLLTQLAPEVRSGNIQNLGVIRVLDYTCNTIGEKQKVLIIITKLEVVFKALDSEIKCEA 120

QY 213 -----TFLVPPTVESKTKQFALPILKISEMNSRFPVVDVM---YTARNENQVQPO 262
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 EKQEEKPAILLSPKEES-----VLSKPTNAP-PLPPVVLKPKQEVKSASQIVNEQ 170

QY 263 NGRVTLDGELGTT---PLLAVNICFKKEGVIK-----NGDVSRYR-----MDMEI 306
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 171 RGNAPAAPARLAWTRRVHPLISLN--PYQGNWIIKVRVTSKGNLRTYKNARGEGCVFNVEL 228

QY 307 TNTDGTPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPT 366
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 TDVGTQIQAT-----MFNEAAKTFYPM 251

QY 367 ISSSEIYLTSPNLRCTNPQ 386
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 252 FELGKVVYISKSLRVANKQ 271

RESULT 8
US-10-371-558-5
; Sequence 5, Application US/10371558
; Publication No. US20030163840A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Replication Protein A and Use
; FILE REFERENCE: 0899D2
; CURRENT APPLICATION NUMBER: US/10/371,558
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/396,149
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/123,896
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/100,690
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-371-558-5

Query Match
Best Local Similarity 3.9%; Score 114; DB 14; Length 630;
Matches 67; Conservative 42; Mismatches 99; Indels 112; Gaps 16;

QY 123 AIPPG---FPYENLSPSQIT--MCPHVLDVRLQLEPFLFLLPMDIWNPF--FHYNQNDP-K 175
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8 SVTPGAVAFVLENASPDAAATGVPVEIVLQVVDLKP-----IGTRFTFLASDGDKIK 60

QY 176 LRLVAMLYTPLRANNSGD---DVFTVSCR-----VLTGPSDFEF----- 212
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 TMLLTQLAPEVRSGNIQNLGVIRVLDYTCNTIGEKQKVLIIITKLEVVFKALDSEIKCEA 120

QY 213 -----TFLVPPTVESKTKQFALPILKISEMNSRFPVVDVM---YTARNENQVQPO 262
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 EKQEEKPAILLSPKEES-----VLSKPTNAP-PLPPVVLKPKQEVKSASQIVNEQ 170

QY 263 NGRVTLDGELGTT---PLLAVNICFKKEGVIK-----NGDVSRYR-----MDMEI 306
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 171 RGNAPAAPARLAWTRRVHPLISLN--PYQGNWIIKVRVTSKGNLRTYKNARGEGCVFNVEL 228

QY 307 TNTDGTPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPT 366
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 229 TDVDTGQIQAT-----MFNEAAKKFYPM 251
QY 367 ISSSEIYLTSPNLRCTNPQ 386
Db 252 FELGKVVYISKGLRVANKQ 271

RESULT 9

US-10-375-553-5
; Sequence 5, Application US/10375553
; Publication No. US20030174240A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Replication Protein A and Use
; FILE REFERENCE: 0899D3
; CURRENT APPLICATION NUMBER: US/10/375,553
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/396,149
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/123,896
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/100,690
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-375-553-5

Query Match 3.9%; Score 114; DB 14; Length 630;
Best Local Similarity 20.9%; Pred. No. 0.16;
Matches 67; Conservative 42; Mismatches 99; Indels 112; Gaps 16;

QY 123 AIPPG---PPYENLSPSQIT--MCPHVIVDVRQLPFLPMPDIWNPF-FHYNQGNDP-K 175
Db 8 SVTPGAVAFVLENASPDATGVPVPEIVLQVVDLKP-----IGTRFTFLASDGKDKIK 60
QY 176 LRLVAMLYTPLRANNSGD---DVFTVSCR-----VLTKPSDPDEF----- 212
Db 61 TWMLTQLAPEVRSGNIQNLGVIRVLDYTCNTTIGEKQKVLIITKLEVVFKALDSEIKCEA 120
QY 213 -----TFLVPPTVESKTKQFALPILKISEMTNSRFPVPVDVM---YTARNENQVVPQ 262
Db 121 EKQEEKPAILLSPKEES-----VVLKSKPTNAP-PLPPVVLKPKQEVKSASQIVNEQ 170
QY 263 NGRVTLDBELLGTT---PLLAVNICKFKGEVIKAK-----NGDVSRYR-----MDMEI 306
Db 171 RGNAAAPAAARLAMTRRVHPLISLN--PYQGNWIIKVRVTSKGNLRYTKNARGGCVFNVEL 228
QY 307 TWTGDTPTDPTDTPGIGSPDFQGLFGVASQRNKEQNAPTRAHEAINTGGHLCQP 366
Db 229 TDVDTGQIQAT-----MFNEAAKKFYPM 251
QY 367 ISSSEIYLTSPNLRCTNPQ 386
Db 252 FELGKVVYISKGLRVANKQ 271

RESULT 10

US-10-372-553-5
; Sequence 5, Application US/10372553
; Publication No. US20040098769A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Replication Protein A and Use
; FILE REFERENCE: 0899D3
; CURRENT APPLICATION NUMBER: US/10/372,553
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/396,149
; PRIOR FILING DATE: 1999-12-15

; PRIOR APPLICATION NUMBER: US 60/123,896
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/100,690
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-372-553-5

Query Match 3.9%; Score 114; DB 16; Length 630;
Best Local Similarity 20.9%; Pred. No. 0.16;
Matches 67; Conservative 42; Mismatches 99; Indels 112; Gaps 16;

QY 123 AIPPG---PPYENLSPSQIT--MCPHVIVDVRQLPFLPMPDIWNPF-FHYNQGNDP-K 175
Db 8 SVTPGAVAFVLENASPDATGVPVPEIVLQVVDLKP-----IGTRFTFLASDGKDKIK 60
QY 176 LRLVAMLYTPLRANNSGD---DVFTVSCR-----VLTKPSDPDEF----- 212
Db 61 TWMLTQLAPEVRSGNIQNLGVIRVLDYTCNTTIGEKQKVLIITKLEVVFKALDSEIKCEA 120
QY 213 -----TFLVPPTVESKTKQFALPILKISEMTNSRFPVPVDVM---YTARNENQVVPQ 262
Db 121 EKQEEKPAILLSPKEES-----VVLKSKPTNAP-PLPPVVLKPKQEVKSASQIVNEQ 170
QY 263 NGRVTLDBELLGTT---PLLAVNICKFKGEVIKAK-----NGDVSRYR-----MDMEI 306
Db 171 RGNAAAPAAARLAMTRRVHPLISLN--PYQGNWIIKVRVTSKGNLRYTKNARGGCVFNVEL 228
QY 307 TWTGDTPTDPTDTPGIGSPDFQGLFGVASQRNKEQNAPTRAHEAINTGGHLCQP 366
Db 229 TDVDTGQIQAT-----MFNEAAKKFYPM 251
QY 367 ISSSEIYLTSPNLRCTNPQ 386
Db 252 FELGKVVYISKGLRVANKQ 271

RESULT 11

US-10-282-122A-66008
; Sequence 66008, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Lianguo
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578

```

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 66008
; LENGTH: 861
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-10-282-122A-66008

Query Match      3.7%; Score 108.5; DB 12; Length 861;
Best Local Similarity 19.3%; Pred. No. 0.86;
Matches 117; Conservative 77; Mismatches 238; Indels 173; Gaps 26;

QY 1 MKMASNDAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVQQNIIDPWIRNPFVQAP 60
Db 46 VELLARVPVPCVDDAAKVFASF-----LAAVAGSASSP-----LISPEYATELLGT 93

QY 61 AGEFTVSPRNPGEIILLDELGLDLPNYPYLAHLARMYNGHAGMEVOIVLAGNAFTAGKII 120
Db 94 LGGYNI---HALIELLDDDKLAPAAKGLKHTLLMFDSDHVOEK---AEKGNKYAQEVILQ 148

QY 121 FAAPGPGFPYENLSPQITMCPHVIIIDVRQLEPFLPMPDIDWNNPFPHYQNDPKRLVA 180
Db 149 SWADAEMFASRAKVEKITVTFK-VDGETNTDLSAPADWSR-----PDIFLHALA 200

QY 181 MLXTP-----LRANNS-----GDDVFTVSCR-----VLTKPSPD 209
Db 201 MLKNPRDGIPTDKPEGVGPIKLEELKAKGHPVAYVGVVGTGSSRSKATNSVIWHTGED 260

QY 210 FEFTFLVPPTVESKTKQFA-----LPILKISEMTNSRFPVPVDMVYTARNENQVVP 261
Db 261 IPF---VP-----NKRFGVCLGGKIAPIFFNTQEDSGALPIEDVVSALKMGDVVDILP 311

QY 262 QNGRVTLDELGLGTTPLLA-----VNIKFKGVIKAKNGDVSRYRMDMEITNT 309
Db 312 YEGKIVKNGETVAEFELKSVLLDEVQAGGRINLIIGRG-LTAKAREA-----LKLPASTA 366

QY 310 DGTPIIDTETDPG-----PIGSPDFQGLIFG-----VASQRNKNQNPATR--- 350
Db 367 FRLPQAPAESKAGFTLAQKMGVGRACGLPEGQGVRRPGTYCEPRMTTVGSQDTGPMTRDEL 426

QY 351 AHEAIINTGGHLCFQISSSEIYLTSPNILECTNPQLPQ-SGLRGITILIRSDNGHCH-- 407
Db 427 KDLACLGSADVMVQSFCHTAAY---PKPVDVTKHKELPAFISTRGVSLRPGDGVJHSW 483

QY 408 -----DMVCTSTPTTPQWRRCRSGNSCCSGHRYPPVPMNRVTWIVLSHKSGFS 460
Db 484 LNRLLLPDTVTGDSHT-----RPIGISFPAGSLV-----APA 519

QY 461 TSTRKLQPLNLRWLPIRFPNDTGRVLFPEARLH-----KQGITVAHT 503
Db 520 AATGWMP-LDM-----PESVLRPSGKLPQGVTLRDLVNAIPLYAIKQGLITVAKA 569

QY 504 GDNFI 508
Db 570 GKQNI 574

```

RESULT 12
 US-10-229-148B-4
 ; Sequence 4, Application US/10229148B
 ; Publication No. US20040091975A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Meiji Seika Kaisha, Ltd.

```

; TITLE OF INVENTION: Midecamycin biosynthetic genes
; FILE REFERENCE: 138451 US
; CURRENT APPLICATION NUMBER: US/10/229,148B
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 210516/2002
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 3696
; TYPE: PRT
; ORGANISM: Streptomyces mycarofaciens
US-10-229-148B-4

Query Match      3.7%; Score 108; DB 16; Length 3696;
Best Local Similarity 19.7%; Pred. No. 8.6;
Matches 147; Conservative 72; Mismatches 223; Indels 306; Gaps 37;

QY 8 AAPSSDGAAG-----LVPEINNEVMPLEPVAGASLATPVVQQNI-IDPWIRNPFVQAPA 61
Db 1930 AGISSFGISGNTNVHIVVEE-----PPEPTAPESLWPDAAAAGDGVWSEEMWRE--VTVP- 1980

QY 62 GEFTVSPRNPGEIILLDELGLDLPNYPYLAHLARMYNGHAG-----LISPEYATELLGT 101
Db 1981 --LMSAHNEAALCDQARRLRADLLAHLAHPADVGLSLITTRTFEHRAAVNGENTEL 2038

QY 102 -----GMEVOIVLAGNAFTAGKIIFAAIPGFPY-----ENLSPSQ 137
Db 2039 IALDDLLLEGPHPLVMRGTAGTADQVVFVPGGSSQWAEANGDLGFRSSVFRETAAHCD 2098

QY 138 ITMCPHV---TIDVRQLEP-----FLLPMPDIWNNF-----FHVNQ 170
Db 2099 AALREYDWSVLDVLRREPDPAPSLDRVDVQVPLFTMMVSLAATWRSLGVEPAAVVGHSQ 2158

QY 171 GN-----DPKRLVAMLYTLR-----ANNSGDDVFTVSCRVLTKPSPDPE 211
Db 2159 GEIAAAHVAGLSLDDAARIIVA-----LRSQAWLQLAGKGWVAVTMSEREL---RPLRE 2210

QY 212 F-----TFLLVPTVESKTKQFALPILKISEMTNSRFPV-PVDVMYTARNENQVVPQ 262
Db 2211 FWGDRLAVAANVSPETCAVAGDPDALAEL-VAEALASQGVPARPIPGVDTAGHSQV--- 2265

QY 263 NGRVTLDELGLGTTPLLAWNICKFKGEVIA-----KNGDVSRYRMDMEITNDGTIDPTE 318
Db 2266 ---DTLEDQL-----REVLPVAPSSSDIPFY-----STVTGGGLDTEAE 2301

QY 319 DTFGPIGSPDFQGLIFGVASQRNKNQNPATRAHEAIINTGGHLCFQISSSEIYLTSPN 378
Db 2302 -----LDADYWRNMRPEPVEFEKATALLADGHD-----VFLE--- 2334

QY 379 ILRCTNPQPLQSGLRGTILIRSDNGHCHDMVGT----- 412
Db 2335 ----TSPHMLAISLQETI---SDAGASAAVLTLRRGGQPRWLGAVACRAYTHGVLEID 2387

QY 413 ---SPTT---PTWPOWR-----CSRGSNCCSSG-----H 437
Db 2388 ABALFGDPRPVGLPTYPFQRYERYWYSPVSRGDDPASLGLDADHPLGGGVGVELGSGDQ 2447

QY 438 RYPVPVVMNRVTWIV-----LSHKSGFSTSTRKLQPLNLRWLPIRFIN 480
Db 2448 MYTARIGTORVPLVDHALMGTVLLPGAVFTDLDALWAGRQTGTGRIEELTLAAPLVL--- 2504

QY 481 PDTGRVLF-----EAR-----LHKQGITVAHTGDNP---IVMPP 512
Db 2505 PESGGVWLRNLNVGAPDTEARRFAVHARPEGAADWTLHAEGLLTAHAADAPDASAVTFS 2564

QY 513 NGYFR-----FEAWNQYSLAPVGTG 534
Db 2565 HGAEQLDGTDFYERFTELGYSGYGFPRG 2592

```

RESULT 13
 US-10-258-951-74

Sequence 74, Application US/10258951
Publication No. US20040033504A1
GENERAL INFORMATION:
APPLICANT: Agarwal, Pankaj
APPLICANT: Murdock, Paul R.
APPLICANT: Rizvi, Safia K.
APPLICANT: Smith, Randall F.
APPLICANT: Xiang, Zhaoying
APPLICANT: Kabnick, Karen
APPLICANT: Lai, Ying-Ta
APPLICANT: Xie, Qing
TITLE OF INVENTION: NOVEL COMPOUNDS
FILE REFERENCE: GP50025
CURRENT APPLICATION NUMBER: US/10/258,951
CURRENT FILING DATE: 2002-10-28
PRIOR APPLICATION NUMBER: PCT/US01/13360
PRIOR FILING DATE: 2001-04-26
PRIOR APPLICATION NUMBER: 60/199,963
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/203,336
PRIOR FILING DATE: 2000-05-11
PRIOR APPLICATION NUMBER: 60/207,087
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/207,546
PRIOR FILING DATE: 2000-05-26
NUMBER OF SEQ ID NOS: 78
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 74
LENGTH: 607
TYPE: PRT
ORGANISM: Homo sapiens
US-10-258-951-74

Query Match 3.7%; Score 107; DB 12; Length 607;
Best Local Similarity 19.4%; Pred. No. 0.7;
Matches 116; Conservative 73; Mismatches 208; Indels 200; Gaps 29;
QY 51 WIRNFWQAPAGFTVSPRNSPGEILL-----DLELGPDLNPLYLAHARMYNGHAGGMEV 105
DB 46 WYWNQF-----FVIEYAGPEPVLIKLSHDVDRGERTKYL-----LTGEGAGTVFV 93
QY 106 QIVLAGNAFTAGKIIFAAITPPGPPYENLSPSQTMCPHVI--IDVRQLEP--FLLPMPD 160
DB 94 IDEATGNHVTKSL-----DREEKAQVLLAQAQVADRASNRPLEPPSEFIKQVD 142
QY 161 IWNN-----PFHYNQGNPKLRLVAMLYTPLRANNSGDDVFTVSCRVLTKPSDPFEPT 213
DB 143 INDNPIFPLGPHY---ATVPENSVGTSVIQTADHADDPSYGSAKLVYTVLDGLPFF 199
QY 214 FLVP-----PTVESKTKQFALPKISEMTNS-----RPFVPVDMYMTARNENQVVQ 260
DB 200 SVDPPQGVVRTAIPNMDRETQBEFLVVIQAKDMGHHMGLSGSTTVTLSDVNDNPPKF 259
QY 261 PQN-----GRVTLDEL-----LGTTPLLAVNICFKPKG-EVIAKNGDV----- 297
DB 260 POSLYQFSVETAGPGLVRLRAQDPDLGDNALMAYSLDGESEAFSISTDLOGRDL 319
QY 298 -----RSYRMDMEITNDGTPTDP-----TEDTPGPIGSD 328
DB 320 LTVRKPLDPESQSRYSFRVEATN---TLIDPAYLRGPPKDVASVRVAVQDAPEP---PA 373
QY 329 FQGLIFGVASQORNKNEQNATRAHEAIINTGGDHLCPQISSEIYLTSPNILRCTNPQPL 388
DB 374 FTQAAVHLTVPENK---AFGT-----LVGQISAADL-----DS 403
QY 389 POSGLRGTLIRSDNGHC-----HDMVCTS-----PTPTTW-- 419
DB 404 PASPIRYSILPHSDPERCFESIQPEEGTIHTAAPLDREARAWNLVLA TELVYPTPAYAS 463
QY 420 --PQWRRCRSRGN-CSSGH-RYPVPVVMNRVTM-----IVLSHSGFSTSTRKL---PQ 468
DB 464 GAPPPLHTAYENCPCICGYLNVSKAYMNVMMAMVLFVFAEHKGGGRGPRQAVDQK 523

QY 469 LNLRWPLIRFIN-----PDTGRVLFEARL-----HKQGITVAHTGDNPIV 509
DB 524 QSTRWELSONCNYLPKSRGVHPGTSMRAQASQLOGSRGTHRNCTRIACHTRVNPIL 580
RESULT 14
US-10-296-115-1211
Sequence 1211, Application US/10296115
Publication No. US20040053248A1
GENERAL INFORMATION:
APPLICANT: Hyseq Inc
TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
FILE REFERENCE: 784PCT
CURRENT APPLICATION NUMBER: US/10/296,115
CURRENT FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: US09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US09/552,317
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 1478
SEQ ID NO 1211
LENGTH: 767
TYPE: PRT
ORGANISM: Homo sapiens
US-10-296-115-1211

Query Match 3.7%; Score 106.5; DB 12; Length 767;
Best Local Similarity 20.0%; Pred. No. 1.1;
Matches 112; Conservative 65; Mismatches 200; Indels 183; Gaps 30;
QY 51 WIRNFWQAPAG-EFTVSPRNSPGEILLDLLELGPDLNPLY--LAHARMYNGHAGGMEVQI 107
DB 204 WKDNEVLTNADMENFVHVHAQDGEVM-----EYTTILHLRQVTFGHEG--RYQC 251
QY 108 VLA---GNAFTAGKIIFAAIPGPPYENLSPSQITM-----C-----BHVLIID- 147
DB 252 VITNHFSGTYSKARLTNVNLPF---TKTPHDITIRTTMARLECAATGHNPQIAWOK 308
QY 148 -----VRQLEPEFLPMPDIWNNPFHYNQGNPKLRLVAMLYTPLRANNSGDDVFTVS 199
DB 309 DGGTDFPAARERRMHVMPDDV---FFITDVKIDD-----AGVYCTAQSAGSISANAT 360
QY 200 CRVLTKPS---PDFEFTFLVPTVESKTKQFALPKISEMTNSRFPVDPVDMYMTARNEN 256
DB 361 LTVLETPSLVWPLEDRVVSVGETVALQCKATGNPPRITWPKGDR-PLSLTERHHLTPDN 419
QY 257 QVVOQPN-----GRVTLD-CELLGT-----TPLLAVNICFKGEVIAKNG- 295
DB 420 QLLVVQNVVAEDAGRYTCMSENTLGTERRAHSQLSVLPAAGCKOGTTV---GIFTIAVVS 476
QY 296 -----DVSRYRMDMEITNDGTPTDPTEDTPGPIGSPDFQGLIFGVASQRN 341
DB 477 SILVTSLVWVCIIYQTRKKSEYSVTNDETVPVP--DVPSYLS--QGTLSDROETVV 531
QY 342 KNEQNATRAHEAIINTGGDHLCPQISSEIYLTSPNI--LRCTNPQ----- 386
DB 532 RTEGGPQANGH---IESNG--VCPRDAS---HFPEPDTHSVACRQPKLCAGSAVHKKPKW 583
QY 387 -----PLQSGLRGTILIRSDNGHCHDMV-----GTSPTTPTWQ 421
DB 584 AMEKAEGTGPCHKMEHGRVVCSD---CNTEVDVCSYRGOAFHPQVPSRDSQAQPSAPNGE 640
QY 422 -----QWRRCSRGNCCSSGHRYFPVVMNRVTWLVLSHSGFSTSTRKLQPL 469
DB 641 PGGSQDEHSPHHQCSRTAAGSCPECQCSLYP-----SNHDEMLTAVKCKPMA 687
QY 470 NL-----RWPLIRFINPDT 483
DB 688 SLDGKGDSSWTLARLYHPDS 707

RESULT 15
US-10-282-122A-65330

[illegible]

```
; Sequence 65330, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 65330
; LENGTH: 861
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
; US-10-282-122A-65330

Query Match      3.7%; Score 106.5; DB 12; Length 861;
Best Local Similarity 19.2%; Pred. No. 1.3;
Matches 116; Conservative 78; Mismatches 238; Indels 173; Gaps 26

Qy      1 MKWAGNDAAPSDGAGLVPEINNEVMPLEPVAGASLATPVGGQQNIIDPWIRNNFVQAP 60
        :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      46 VELLAHRVPPGVDDAKVKASF-----LAAVAEGSASSP-----LVSPKYATELLGTGM 93

Qy      61 AGEFTVSPRNSPGEILLDELPDNLNVLAHARMYNGHAGGMEVQIVLAGNAFTAAGKII 120
        :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      94 LGGYNF---HALIEILDDDKLAPIAAKGKHTLLMFDSFHDVOEK--AEKGNYAOEVLIQ 148

Qy      121 FRAIPGFPVENLPSQTMTCPHFVIIDVRQLEPLLPMPIWNPFHYNGNDPKRLVA 180
        :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      149 SWADAWEFASRAKVKEKITTVFK-VDETWTDDLSPADWSR-----PDIFLHALA 200

Qy      181 MLVTP-----LRANNS-----GDVFVTSCR-----VLTKPSPD 209
        :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      201 MLKNRPDGITDPKPGEVGFIKLLELKAKGHPAVAVGVGVGTSSRSKSATNSIWMTHGED 260

Qy      210 FEFTFLVPPTVESKTYQFA-----LPIIKISEMNTSRPVPDVVMYTARNENOVQP 261
        :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      261 IPF---VP-----NKRFGGVCLOGKIAPIFNTQEDSGALPIEDVDVSALKMGDVVNILP 311

Qy      262 QNGRVTLDELGLTGTTPLA-----VNICKFKGEVIANKGDRVSYRMDMBITNT 309
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:36:02 ; Search time 10.4388 Seconds
(without alignments)
4985.230 Million cell updates/sec

Title: US-09-926-799-11

Perfect score: 2894

Sequence: 1 MKMASNDRAAPSSDGAAGLVLP.....VNQFYSLAPVGTGKRRRVQ 541

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: Piri:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1620.5	56.0	542	2 S60616	capsid protein - h
2	1562	54.0	539	2 S40111	capsid protein - h
3	1004.5	34.7	530	2 B37471	capsid protein - N
4	966	33.4	546	2 B37491	major capsid prote
5	317	11.0	2344	1 RRRWRH	genome polyprotein
6	304	10.5	2344	2 S53399	genome polyprotein
7	304	10.5	2344	2 S64740	genome polyprotein
8	279.5	9.7	576	2 A53982	capsid protein - E
9	271.5	9.4	703	1 C48562	coat protein - San
10	262.5	9.1	671	1 VCVWF9	coat protein - fel
11	262	9.1	702	1 A48562	coat protein - San
12	260	9.0	668	1 VCVWFF	coat protein - fel
13	254	8.8	668	1 JQW2F	capsid protein - f
14	248.5	8.6	668	2 JQ2354	capsid protein - f
15	248.5	8.6	668	2 JQ2356	capsid protein - f
16	126.5	4.4	2194	1 GNNYF7	genome polyprotein
17	126	4.4	3085	2 T00327	polyprotein - infe
18	118.5	4.1	2333	1 GNNY2F	genome polyprotein
19	117.5	4.1	2205	1 GNNY2W	genome polyprotein
20	117.5	4.1	2303	2 S13554	genome polyprotein
21	117	4.0	2175	1 GNNYBE	genome polyprotein
22	115.5	4.0	2207	1 GNNY1P	genome polyprotein
23	115.5	4.0	2209	1 GNNY2P	genome polyprotein
24	114	3.9	630	2 T03582	probable replicat
25	113.5	3.9	2206	1 GNNY21	genome polyprotein
26	112.5	3.9	2303	1 GNNYTP	genome polyprotein
27	112	3.9	2185	1 GNNYSH	genome polyprotein
28	112	3.9	2301	1 GNNYTN	genome polyprotein
29	112	3.9	2332	1 GNNY4F	genome polyprotein

30	111	3.8	2183	1 GNNYB4	genome polyprotein
31	110	3.8	2209	1 GNNY3P	genome polyprotein
32	109	3.8	929	2 A44048	genome polyprotein
33	109	3.8	2185	1 GNNYSV	genome polyprotein
34	109	3.8	2290	1 GNNYE	genome polyprotein
35	108.5	3.7	861	2 A81801	aconitate hydratase
36	108.5	3.7	2207	1 GNNYSP	genome polyprotein
37	108	3.7	2206	2 S03822	genome polyprotein
38	108	3.7	13055	2 T16580	hypothetical prote
39	106.5	3.7	940	2 D89723	protein F39D8.1b (
40	106.5	3.7	945	2 T21998	hypothetical prote
41	106.5	3.7	2481	2 A43908	fibronectin - Afri
42	106	3.7	2201	1 GNNYA9	genome polyprotein
43	106	3.7	2206	1 GNNY4P	genome polyprotein
44	106	3.7	2207	2 S09553	genome polyprotein
45	104.5	3.6	2303	1 GNNYTM	genome polyprotein

ALIGNMENTS

RESULT 1

S60616

capsid protein - human calicivirus (strain Melksham)

C:Species: human calicivirus

A:Variety: strain Melksham

C:Date: 23-May-1997 #sequence_revision 23-May-1997 #text_change 28-Jul-2000

C:Accession: S60616

R:Green, S.M.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.

Virus Res. 37, 271-283, 1995

A:Title: Capsid diversity in small round-structured viruses: molecular characterization

A:Reference number: S60615; MUID:96136658; PMID:8533462

A:Accession: S60616

A:Molecule type: genomic RNA

A:Residues: 1-542<GRE>

A:Cross-references: EMBL:X81879; NID:G976077; PIDN:CAA57462.1; PID:G976079

A:Experimental source: strain Melksham

A:Note: it is uncertain whether Met-1 or Met-3 is the initiator

C:Superfamily: human calicivirus capsid protein

C:Keywords: capsid protein; coat protein

Query Match	56.0%	Score	1620.5	DB	2	Length	542
Best Local Similarity	58.1%	Pred. No.	2e-116				
Matches	324	Conservative	68	Mismatches	133	Indels	33
Gaps	7						
Qy	1	MMASNDRAAPSSDGAAGLVLP	INNEVMPLEPVAGASLATPVVQQNIIDPWIRNNFVQAP	60			
Db	1	MMASNDRAAPSSDGAAGLVLP	ESNEVMALEPVAGAAAPVTGQTNIIDPWIRAFVQAP	60			
Qy	61	AGEFTVSPRNSPGEILLDLELGPDLNPLYLAHLMRYNGHAGMEVQIVLAGNAFTAGKII	120				
Db	61	NGEFTVSPRNPAPGVEVLLNLELGPDLNPLYLAHLMRYNGVAGMEVQVVLGNAGNAFTAGKIV	120				
Qy	121	FAATPGPFPYENLSQITMCPHVIIDVROLEPPLPMPDIWNFFHNGNDPKRLVA	180				
Db	121	FAAVPFPFPYENLSQITMCPHVIIDVROLEPPLPMPDIWNFFHNGNDPKRLVA	180				
Qy	181	MLYTPLRANNSGDDVFTVSCRVLTTPKPSDFEFTFLVPPTVESKTKQFALPILKISEMTNS	240				
Db	181	MLYTPLRANNSGDDVFTVSCRVLTTPKPSDFEFTFLVPPTVESKTKQFALPILKISEMTNS	240				
Qy	241	RFPVVDVVMYTARNENQVVPQNGRVTLLDGLLGTTPLLAVNICKFKGEVIAKNGDVSRY	300				
Db	241	RFPVVDVVMYTARNENQVVPQNGRVTLLDGLLGTTPLLAVNICKFKGEVIAKNGDVSRY	300				
Qy	301	RMDMEITWDGTPIDPTDTPGPGSPDFQGLIFGVASQRNK-----NQONPATRAHEAII	356				
Db	300	LNNVTITNLNGSPDFPDSIDIAPLGVDPFQGRVFGVVISQRDKQNAAGHSEPPANRGHDVAV	359				
Qy	357	NTGGDHLCPQISSSEIYLTSPNILLRCTNTPQLPQSGLAGTILIRSDNGHCHDMV-----	410				
Db	360	PTYTAQYTPKLGQIQIGTQWTDLLTVPNQVFPVFTPVGL-----NDTEHFENQVMPRVVAG	412				

```
QY 411 -----GTSPT-TPTWPOWRRCRSGNCCSGHRRPVPVVMNRVTWVLSHKSGFSTST 463
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 413 ALNLTNLTAPSVAPVPPGRLFFRSHLPKGGYGNPAIDCLLPQEWQHFYQEA----- 467
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 464 RXLPQLNLWRPLIRFINDPTGRVLPEARLHKOGFTITVAHTGDNPIVMPDNGYFRFEAWN 523
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 468 --APGMS-EVALVYINPTGRALFEAKUHRAGFMVSSNTSAPVVVPANGVFRDSSWN 524
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 524 QFYSLAPVGTGKRRVQ 541
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 525 QFYSLAPMGNGRRVQ 542
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 2
S40111
capsid protein - human calicivirus (strain Bristol isolate B493)
C:Species: human calicivirus
A:Variety: strain Bristol isolate B493
C:Date: 25-Dec-1994 #sequence_revision 27-Feb-1997 #text_change 28-Jul-2000
C:Accession: S40111
R:Green, S.M.; Dingle, K.E.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
submitted to the EMBL Data Library, December 1993
A:Description: Human enteric caliciviridae: a new prevalent SRSV group defined by RNA-de
A:Reference number: S40111
A:Accession: S40111
A:Molecule type: genomic RNA
A:Residues: 1-539 <GRE>
A:Cross-references: EMBL:X76716; NID:G436410; PIDN:CAA54134.1; PID:G436411
A:Experimental source: human enteric calicivirus strain Bristol isolate B493
C:Superfamily: human calicivirus capsid protein
C:Keywords: capsid protein; coat protein

Query Match 54.0%; Score 1562; DB 2; Length 539;
Best Local Similarity 54.3%; Pred. No. 6e-112;
Matches 305; Conservative 76; Mismatches 133; Indels 48; Gaps 8;

QY 1 MKWASNDAASSDGAAGLVPEINNEVMPLEPVAGASLATPVVQQNIIDPWIRNFFVQAP 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MKWASNDANPSGSAANLVPEVNEVNALEFPVVGAAIAAPVACQNVIDPWIRNFFVQAP 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 AGEFTVSPRNSPGEILLDLBELGPDLPNLYLAHARMYNGHAGGMEVQIVLAGNAFTAGKII 120
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 GGEFTVSPRAPGEILLASAPLGPDLPNLYLSRMVNGYAGGVEQVIVLAGNAFTAGKVI 120
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 FAIIPGFPVENLSPQITWCPHVIIIDVRLQLEPFLPMPDIWNNFHYNQNDPKRLVA 180
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 FAAPVPPNFTGLSPSQVTFPHIIVDVRQLPEVLIPLDVRNMFHYNQANDSTKLIA 180
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 MLYTLRANNSGDDVFTVSCRVLTKPSDFEFTFLVPTVESKTKQFALPILKISEMTNS 240
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 MLYTLRANNSGDDVFTVSCRVLTKPSDFDFFLVPPTVESKTKPTVPLVTEMSNS 240
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 241 RFPVPVDMVTARNENQVQPNQGRVTLDBELLGTTPELLAVNICKFKEGVIANKGDVRSY 300
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 RFPPLKLYITGPSSAFVQVQNGRCITDGVLLGTTQLSAVNICNFRGVDTHIAG---SH 297
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 301 RMDMEITNDGTPIDPTEDTPGTPGDFQGLFPGVASQRNKNQONPATRAHEALINTGG 360
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 298 DYTWNLASQWNSNDYDTEEPAPLGTDFVGKIQLLTQTTR--ADGSTRAHKATVSTGS 355
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 361 DHLCPQISSSIYLTSPNLRCTNPQLPQSGLRGTILIRSDNGHCHDMVGTSTPTTWP 420
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 356 VHFTPKLSGVQFTDNDFOAGQNTKFTPVG-----VIQGDHQN-----P 399
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 421 QWRRCRSGNCCSGHRRP--VPVVMNRVTWVLSHKSGFSTSTRKLPLQNL-----R 472
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 400 QWLLPNYSGRTGHNVLAPAVPTFGEQLLFFRSTMPGCS-----GYPNMNLCLLPQE 455
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 473 W-----PLIRFINDPTGRVLPEARLHKOGFTITVAHTGDNPIVMPDNGYFR 517
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 456 WTLHFYQBAAPQAQSDVALLRFVNDPTGRVLFECKLHSGYITVAHTGIDVLVLPNGYFR 515
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 518 FEAWNQFYSLAPVGTGKRRR 539
```

```
Db 516 FDSWVNFYTLAPMGNGTGRRR 537
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 3
B37471
capsid protein - Norwalk virus
C:Species: Norwalk virus
C:Date: 24-Feb-1994 #sequence_revision 25-Apr-1997 #text_change 28-Jul-2000
C:Accession: B37471
R:Jiang, X.; Wang, M.; Wang, K.; Estes, M.K.
Virology 195, 51-61, 1993
A:Title: Sequence and genomic organization of Norwalk virus.
A:Reference number: A37471; MUID:93303939; PMID:8391187
A:Accession: B37471
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: genomic RNA
A:Residues: 1-530 <JTA>
A:Cross-references: GB:M87661; NID:G1061311; PIDN:AAB50466.1; PID:G1061313
A:Note: sequence extracted from NCBI backbone (NCBIP:134157)
C:Superfamily: human calicivirus capsid protein

Query Match 34.7%; Score 1004.5; DB 2; Length 530;
Best Local Similarity 40.8%; Pred. No. 3.6e-69;
Matches 234; Conservative 73; Mismatches 180; Indels 87; Gaps 14;

QY 1 MKWASNDAAAPSSDGAAG--LVPEIN-NEVMPLEPVAGASLATPVVQQNIIDPWIRNFF 56
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MMWASKDTATSSVDGASGAGQLVPEVNASDPLAMPDVAGSSTAVATAGQVNPIDPWIRNFF 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 57 VOAPAGEFTVSPRNSPGEILLDLBELGPDLPNLYLAHARMYNGHAGGMEVQIVLAGNAFTA 116
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 VOAPQGEFTISPNNTPGDVLFDLSLGLPHNLFLLHLSQMYNGWVGNMVRIMLAGNAFTA 120
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 117 GKIIIPAAITPPGFPVENLSPQITWCPHVIIIDVRLQLEPFLPMPDIWNNFHYNQNDPKL 176
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 GKIIIVSCIPPGFGSHSLTIAQATLFPFHVIAVRLDPILEVEDVRNVLFHNNDRNQOTM 180
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 177 RLVMALYTLRANNSGDDVFTVSCRVLTKPSDFEFTFLVPTVESKTKQFALPILKISE 236
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 RLVCMLYTLRTRGGTGDGDFVAGRVMTCPSPDFNLFVLPVTEQKTRPTLPNPLSS 240
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 237 MTNSRFPVPVDMVTARNENQVQPNQGRVTLDBELLGTTPELLAVNICKFKEGVIANKGD 296
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 LNSRAPLPISSIGISPDNVQSPQNGRCITDGLRGVGTTPVLSHVAKIRG---TSNGT 297
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 297 VRSYRMDMEITNDGTPIDPTEDTPGTPGDFQD-----ILFGVASQRNKNQONPAT 349
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 298 V-----INLTEDGTGPHFPFG-PAPIGFDLGGCDWHINMTQFGHSSQTYD----- 344
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 350 RAHEALINTGGDHLCPQISSSIYLTSPNLRCTNPQLPQSGLRGTILIRSDNGHCHDM 409
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 345 -----VDITPTDTPVPHLGS-----IQANGICSGNY 369
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 410 VG-----TSPTPTTPWQ--QWRRCRSGNCCSGHRRP--VPVVMNRVTWVLSHKSGFS 460
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 370 VGVLSWISPPSPSGSQVDLMKIPNYGSSITEATHLAPSVYPPGFEVLVFFMSKMPGCG 429
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 461 TSTRK--LPQLNL-----RWPLIRFINDPTGRVLPEARLHKOGFTITVAHTG--- 504
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 430 AYNLPCLLQFYISHLASQAPTVGEAALLHYVDPTGRNLGELFRAYPDGFLTCTVPNGAS 489
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 505 DNPVMPDNGYFRFEAWNQFYSLAPVGTGKRR 538
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 490 SGPQLPQINGVVFVFSWVSFRYQLKPVGTASSAR 523
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
B37491
major capsid protein [similarity] - Southampton virus
N:Alternate names: orf2 protein
C:Species: Southampton virus
C:Date: 22-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 18-Aug-2000
```

```

A:Molecule type: genomic RNA
A:Residues: 1-2344 <MBY>
A:Cross-references: GB:M67473
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: nucleotidyltransferase

Query Match      11.0%; Score 317; DB 1; Length 2344;
Best Local Similarity 28.3%; Pred. No. 1.7e-15;
Matches 139; Conservative 55; Mismatches 169; Indels 128; Gaps 28;

QY      8 AAPSSDGA---AGLVPEINNEVMLEPVA-----GASLATPVVG---QQNIID 49
DB      1772 AAPQGEAGTATASVPGTTTDMGCVATTSTVITAENSSASIAIATAGICGPPQQVDDQE 1831

QY      50 PHIRNNFVOAPGEFTVSPRNSPGEILLDELGLDLPNLYLAHLARMYNGHAGMEVQIVL 109
DB      1832 TW-RTNFYNDV--FTWSVADAPGSLYTVQHSPPNNPFTAVLSQMYAGWAGCMQFRFV 1888

QY      110 AGNAFTAGKIIFAATPPGPPYENLSPS-QITWCPHVIIDVROLEPPELLMPDIDWNNFFHY 1688
DB      1889 AGSGVFGGRVRAVIPPGI---EIGFGLVRFQPHVVIDARSLEPVTIITPDLRPNMYH- 1944

QY      169 NQGNdpKL--RLVAMLYTEPLRANSGDDVFTVSCRVLTKPSDPDFETFLVPPTVESKTKQ 226
DB      1945 -PTGDPGLVPTVLVSYNML-INPFGSTSAIQVTETRPSEDFEFWMIRAPS--SKT-- 1998

QY      227 FALPILKISEMTNSRFPVDMYTARNENQVV--QPQNGRVT-----LDGELLG-TT 276
DB      1999 -----VDSISFAGLLITPVLTVGNDRNWNQIVGLQVPFGGFSTCNRHWNLNGSTYGMSS 2054

QY      277 PLLAVNICFKGEVIKAGDVRSYRMDMEITN-----TDGTPIDPTBDTPGPIGSPD- 328
DB      2055 P-----RF-GDIDHRRGSA-SY-SGSNATNVLPQWYANAGSAIDNPISQVAPDGFDDM 2104

QY      329 -----PQG-----ILFGVASQRKNEONPAT-RAHEAIIINTGG-DHLCPQISSS----- 370
DB      2105 SFVPNGPGGIPAAGWVGFGAIWNNSGAPNVTTVQAYELGFATGAPGNLOPTNTTSGAQT 2164

QY      371 ---EYI-----LTSFNILRCT-NPQPLPQSGLRGTILIRSDNGHC 406
DB      2165 VAKSIYAVVTGAQNPAGLFVMASGIISTPNAQRITYTQP----- 2205

QY      407 HDMVGTSPTTP 417
DB      2206 -DRIVTTGCTP 2215

RESULT 6
S55399
genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)
C:Species: rabbit hemorrhagic disease virus
A:Variety: isolate BS89
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 18-Jun-1999
C:Accession: S55399
R:Rossi, C.
submitted to the EMBL Data Library, May 1995
A:Reference number: S55399
A:Accession: S55399
A:Molecule type: genomic RNA
A:Residues: 1-2344 <ROS>
A:Cross-references: EMBL:X87607; NID:g854640; PIDN:CAA60910.1; PID:g854641
A:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: polypeptide

Query Match      10.5%; Score 304; DB 2; Length 2344;
Best Local Similarity 28.9%; Pred. No. 1.7e-14;
Matches 125; Conservative 56; Mismatches 177; Indels 74; Gaps 23;

QY      33 AGASLATPVVG-----QQNIIDPWIRNNFVOAPGSEFTVSPRNSPGEILLDELGLDLPN 88
DB      1811 SSASATATAGIGGPPQVDDQETW-RTNFYNDV--FTWSVADAPGSLYTVQHSPPNNPF 1867

```

```
QY 89 LAHLARMYNGHAGMEVQIVLAGNAFTAGKIIFAAIPGCPFYENLSPS-QITMCPHVLID 147
Db 1868 TAVLSQMYAGWAGMGQFRFIVAGSGVFGGRUVAIVPPGI-----EIGPGLVQRQPHVVID 1924
QY 148 VRQLEPFLPMPDIDWNNFFHYNQGNPKL--RLVAMLYTPLRANNSGDDVFTVSCRVLTK 205
Db 1925 ARSLSEPVITMPDLRPNMYH--PTGDPGLVPTLVLSVNNL-INPFGGSTSAIQVTVETR 1981
QY 206 PSPDFEFTFLVPPTVESKTKQFALPKISEMTNSRFPVVDVMTYARNENQV--QPON 263
Db 1982 PSEDFEFVMIRAPS--SKT-----VDSISPAGLLTTPVLTVGVNDNRWNGQIVGLQVPV 2033
QY 264 GRVT-----LDGELLG-TTPLLAVNICFKFGEVIKNGDVRSYRMDMEITN----- 308
Db 2034 GGFSTCNRHWNLNGSTYGSPPFA-----DIDHRRGSA-SY-PGSNATNVLOFWYA 2083
QY 309 TDGTPIDPTEDTPGPIGSPD-----FOG-----ILFGVASQRNKNQONPAT-RAHEA 354
Db 2084 NAGSAVDNFIQVAPDGFDMGSFVFPNGGIPAAQWVGFGAIWNSGAPNVTTVQAYEL 2143
QY 355 IINTGG-DHLCFQISSEIYLTSPNLRCTNPQPLQSGL-----RGTLIRSDNGHCH-- 407
Db 2144 GFATGAPGNLQPTNTSGQTVAKSIYAVVTGTAQNPAGLFVMAAGVISTPANAITYTP 2203
QY 408 --DMVGTSTPTP 417
Db 2204 QPDRIVTTGTP 2215
RESULT 7
S64740 genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)
N;Contains: VP60 protein
A;Variety: isolate AST/89
C;Date: 12-Jul-1996 #sequence revision 26-Jul-1996 #text_change 20-Jun-2000
C;Accession: S64740; S46944; S49018; S65012
R;Boga, J.; Martin-Alonso, J.; Boga, J.; Parra, F.
submitted to the EMBL Data Library, May 1995
A;Description: Genomic organization of rabbit hemorrhagic disease virus determined by di
A;Reference number: S64740
A;Accession: S64740
A;Molecule type: genomic RNA
A;Residues: 1-2344 <CAS>
A;Cross-references: EMBL:Z49271; NID:g1182032; PIDN:CAA89265.1; PID:g1150552
R;Boga, J.; Casais, R.; Marin, M.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra, F.
submitted to the EMBL Data Library, July 1993
A;Description: Molecular cloning, sequence and expression of the capsid protein gene fr
A;Reference number: S46944
A;Accession: S46944
A;Molecule type: genomic RNA
A;Residues: 1650-2344 <BOG>
A;Cross-references: EMBL:Z24757; NID:g515622; PIDN:CAA80881.1; PID:g515623
R;Parra, F.; Boga, J.A.; Marin, M.S.; Casais, R.
Virus Res. 27, 219-228, 1993
A;Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus supp
A;Reference number: S49018; MUID:93255896; PMID:8488721
A;Accession: S49018
A;Molecule type: genomic RNA
A;Residues: 1650-1796 <PAR>
A;Cross-references: EMBL:Z24757
A;Experimental source: isolate AST/89
A;Accession: S65012
A;Molecule type: protein
A;Residues: 1767-1779;1875-1877,'X',1879-1881;1936-1938,'X',1940-1941 <PAW>
C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C;Keywords: blocked amino end; polyprotein
Query Match 10.5%; Score 304; DB 2; Length 2344;
Best Local Similarity 29.6%; Pred. No. 1.7e-14;
Matches 128; Conservative 55; Mismatches 175; Indels 74; Gaps 24;
```

```
QY 33 AGASLATPVVG---QQNIIDPWIRNNFVQAPAGEFTVSPRNSPGEILLDLGLDPLNYP 88
Db 1811 SSASIAITAGTGGPQQVDDQETW-RTNFYNDV--FTWSVADAPGSLTYTVQHSPPQNP 1867
QY 89 LAHLARMYNGHAGMEVQIVLAGNAFTAGKIIFAAIPGCPFYENLSPS-QITMCPHVLID 147
Db 1868 TAVLSQMYAGWAGMGQFRFIVAGSGVFGGRUVAIVPPGI-----EIGPGLVQRQPHVVID 1924
QY 148 VRQLEPFLPMPDIDWNNFFHYNQGNPKL--RLVAMLYTPLRANNSGDDVFTVSCRVLTK 205
Db 1925 ARSLSEPVITMPDLRPNMYH--PTGDPGLVPTLVLSVNNL-INPFGGSTSAIQVTVETR 1981
QY 206 PSPDFEFTFLVPPTVESKTKQFALPKISEMTNSRFPVVDVMTYARNENQV--QPON 263
Db 1982 PSEDFEFVMIRAPS--SKT-----VDSISPAGLLTTPVLTVGVNDNRWNGQIVGLQVPV 2033
QY 264 GRVT-----LDGELLG-TTPLLAVNICFKFGEVIKNGDVRSYRMDMEITN----- 308
Db 2034 GGFSTCNRHWNLNGSTYGSPPFA-----DIDHRRGSA-SY-PGSNATNVLOFWYA 2083
QY 309 TDGTPIDPTEDTPGPIGSPD-----FOG-----ILFGVASQRNKNQONPAT-RAHEA 354
Db 2084 NAGSAVDNFIQVAPDGFDMGSFVFPNGGIPAAQWVGFGAIWNSGAPNVTTVQAYEL 2143
QY 355 IINTGG-DHLCFQISSEIYLTSPNLRCTNPQPLQSGL-----RGTLIRSDNGHCH-- 407
Db 2144 GFATGAPGNLQPTNTSGQTVAKSIYAVVTGTAQNPAGLFVMAAGVISTPANAITYTP 2203
QY 408 --DMVGTSTPTP 417
Db 2204 QPDRIVTTGTP 2215
RESULT 8
A53982 capsid protein - European brown hare syndrome virus
C;Species: European brown hare syndrome virus
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 28-Jul-2000
C;Accession: A53982
R;Wirblich, C.; Meyers, G.; Ohllinger, V.F.; Capucci, L.; Bakens, U.; Thiel, H.
J. Virol. 68, 5164-5173, 1994
A;Title: European brown hare syndrome virus: relationship to rabbit hemorrhagic disease
A;Reference number: A53982; MUID:94309183; PMID:7518531
A;Accession: A53982
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-576 <WIR>
A;Cross-references: GB:U09199
C;Superfamily: human calicivirus capsid protein
Query Match 9.7%; Score 279.5; DB 2; Length 576;
Best Local Similarity 25.6%; Pred. No. 1.7e-13;
Matches 121; Conservative 60; Mismatches 170; Indels 121; Gaps 23;
```

```
QY 17 GLVPEI--NNEYNMPEFVAGASLATPVVG-----QQNIIDPWIRNNFVQAPAGEFTVSPRN 70
Db 25 GMDPGVAGTDDVTADNVA-ASVATAGIGGPPQASQPSQSRVNFYFN---DVFTWSVTD 80
QY 71 SGEILLDLLELGLDNLNLYLAHARMYNGHAGMEVQIVLAGNAFTAGKIIFAAIPGPPFY 130
Db 81 AFGSILYSVQHSFPQNNPFTQVLSQMYAGWAGMGQFRFIVAGSGIFGGRVLCALIPPGI-- 138
QY 131 ENLSPS-QITMCPHVIIDVRQLEPFLPMPDIDWNNFFHYNQGNQDNPKL--RLVAMLYTPUR 187
Db 139 -QIQFGLVQRQPHVVIDARSLSEPVITMPDLRPNMYH--PTGDPGLVPTLVSVVNNL- 194
QY 188 ANNSGDDVTVCRLVTKSPDFEFTFLVPP---TVESKTKQFAL--PILKISEMTNSRF 242
Db 195 INPFGGTSAIQVTVETRESEDFEFLIRAPSKTVDSVNPSPVLLTTPVL-TCAGSDNRW 253
QY 243 FVPVDVMTYARNENQVQP-----QNGRVITLDGELLGTTPLLAIVNICFKFGEVIKNG 295
```

Db 254 GAPI-----VGLQPVGCGFSTNRHWNWNGSTYGSWSPRFPDIDHPSGNVSYPTG 303

QY 296 DVRSYRMDMEITNT-----DGTPTDPTEDTPGPIGSPDPFOGILF-----GVA 337

Db 304 SA-----TNTIETWYANAGATTNPISNIAPDGFDFDGAIPSPGTTIPTGAWVGFG 354

QY 338 SQNKNEQNP-----ATPAHE-----AIIITGGDHLCP-----QISSSE 371

Db 355 QVWNASNGTFYGVGTVOAYELGFANGAPSSIRPVTITTTGAQLVAKSYGYVAIAQNSSAGI 414

QY 372 IYL-----TSPNLRCT-NPQOPQSGLRGTLIRSDNGHCHDMVGTSTPTP 417

Db 415 IFLSKMVGSTPGVAATYTFQP-----SAIVTTGTP 446

RESULT 9

C48562

coat protein - San Miguel sea lion virus (serotype 4)

N:Alternate names: capsid protein

C:Species: San Miguel sea lion virus

C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999

C:Accession: C48562

R:Neill, J.D.

Virus Res. 24, 211-222, 1992

A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lion virus

A:Reference number: A48562; MUID:92410750; PMID:1529644

A:Accession: C48562

A:Molecule type: Genomic RNA

A:Residues: 1-703 <NEI>

A:Cross-references: GB:M87482; NID:g334886; PIDN:AAAL6220.1; PID:g334888

A>Note: sequence extracted from NCBI backbone (NCBIP:113567)

C:Superfamily: feline calicivirus coat protein

C:Keywords: capsid protein; coat protein; glycoprotein

F:89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.4%; Score 271.5; DB 1; Length 703;

Best Local Similarity 27.0%; Pred. No. 9.5e-13;

Matches 106; Conservative 57; Mismatches 160; Indels 69; Gaps 18;

QY 8 AAPSSDGAAGLPEINNEWPLE---PVAGASLATPVVQQNIIDPWIRNPFVQAPAGE 63

Db 151 AESDGGCSAEIVTEEGTVVQQOPAPAPTALATASTG-KSVEQEWMTFFSYHTSINW 209

QY 64 FTVSPRNSPGEILLDLELGLDLPYLAHLARMYNGHAGMEVQIVLAGNAFTAGKIIIFAA 123

Db 210 STV---ESQKILYQALNPSINPYLDHIAKLYSTWSGIDVRFVSGSGVFGKLAALL 266

QY 124 IPPGF-PYENLSPSQITMCPHVIIDVRQLEPFLPMPDIWNNPFHYNQNDPKLRVAML 182

Db 267 VPPGVEPIESVSMLOQ---PHVLFARQTEPVIFTIPDIRKTLFHSMDTD-TTKLVINP 322

QY 183 YTPLRANSGDDVFTVSCVRLTKSPDPFEFLVLPPTVESKTKQFALPILKISEMTNSRF 242

Db 323 Y-----ENGVENKTCITIVETRPSADFTFALLKPP--GSLIKHGSIP-----SDL----- 366

QY 243 PVPDVMYTARN-----ENQVVOPO---NGRVTLDGELLG-TTPLLAVNTCKPKGEVI 291

Db 367 -TPRNSAHMGNRWSTISGFSVQPRVFSQNRHDFDSTTTGWSPTFPYPIKIQGKVG 425

QY 292 AKNGDVSRYMDMEITNTDGTPTDPTEDTPG-PIGSPDPF-----GILFGVASOR 340

Db 426 SNN-----KWFHVIDTDKALV-----FGIPDGWPDTTIPDETAKTNGNFSYGESYRA 472

QY 341 KNNEQNPAFRAHEAIIINTGGDHLCPQISSSEI 372

Db 473 GSTTIKP-----NENSTHFKGTYICGTLSTVEI 500

RESULT 10

VCHWF9

coat protein - feline calicivirus (strain F9)

N:Alternate names: capsid protein

C:Species: feline calicivirus

C>Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 28-Jul-2000

C:Accession: B43382; C45538; PQ0407; S23702

R:Carter, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.; Turner, P.C.

Virol. 190, 443-448, 1992

A:Title: The complete nucleotide sequence of a feline calicivirus.

A:Reference number: A43382; MUID:92410623; PMID:1529544

A:Accession: B43382

A:Molecule type: Genomic RNA

A:Residues: 1-671 <CAR1>

A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879

R:Carter, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.

Arch. Virol. 122, 223-235, 1992

A:Title: Identification and sequence determination of the capsid protein gene of feline calicivirus

A:Reference number: A45538; MUID:92117861; PMID:1731695

A:Accession: C45538

A:Molecule type: Genomic RNA; protein

A:Residues: 1-671 <CAR2>

A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879

A:Experimental source: strain F9

A>Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)

R:Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.

J. Gen. Virol. 73, 2429-2433, 1992

A:Title: The cloning, sequencing and expression of a major antigenic region from the feline calicivirus

A:Reference number: PQ0407; MUID:93019069; PMID:1402818

A:Accession: PQ0407

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 217-266 'SI', 269-336 'V', 338-395 'S', 397-412 'A', 414-521 'EV', 524-671 <GUI>

A:Cross-references: PIDN:AAB23553.1; PID:g257083

C:Superfamily: feline calicivirus coat protein

C:Keywords: capsid protein; coat protein; glycoprotein

F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.1%; Score 262.5; DB 1; Length 671;

Best Local Similarity 25.3%; Pred. No. 4.3e-12;

Matches 113; Conservative 59; Mismatches 167; Indels 107; Gaps 20;

QY 30 EPVAGASLATPVVQQNIIDPWIRNPFVQAPAGEFTV---SPRNSPGEILLDLELGLDIN 86

Db 145 EFSAQMSAADMATKSCVDSEW-----EAPFSPTSVNWSSTEQKILFKQSLGPLN 198

QY 87 PYLAHLARMYNGHAGMEVQIVLAGNAFTAGKIIIFAAIPPGF-PYENLSPSQITMCPHVI 145

Db 199 PYLEHLAKLYVAWSGSIIEVRFISGSGVFGKLAIVVPPGVDPVQSTSMLOQ---PHVL 255

QY 146 IDVRQLEPFLPMPDIWNNPFHYNQNDPKLRVAMLVTPLRANSGDDVFTVSC--RVL 203

Db 256 FPARQVEPVIFCLDLRSLYHLMSDTD-TTSLVIMVNDL-INPYANDANSGCCIVTVE 313

QY 204 TKPSPDFEFTFLVPPPTVESKTKQFALPILKISEMTNSRFP---VP-----VDVMYTARN 254

Db 314 TKPGDPDFKHLKPPG-----SMLTHGSIPLSLIPKTSLSLWIGNRYWSDI 358

QY 255 ENQVVOPO---ONGRVTLDGELLG-TTPL---LAVNICKFKGEVIKNGDVSRYM----- 302

Db 359 TDFVIRPFVQANRHDFENQETAGNSTPRFRPISVTITEQNGAKLG-IGVATDIYVPGIP 417

QY 303 -----DMEITNTDGTPTDPTDTE--DTPGPI-CSPDFQGLFLGVASORKN 343

Db 418 DGMWPTTIPGELIPAGDYAITNGNDITATGYDTADIIKNTNFRGMYICGSLQRAWG 477

QY 344 EQNPATRAHEAIIINTGGD-----HLCPOISSEIYL-----TSPNII----- 380

Db 478 DKKISNTAFITATLDGDNKNKINFCNTIDQSKIIVFQDNHVGKKAQTSDDTLALLGYTG 537

QY 381 -----RCTNPQPLPQSGLRG 395

Db 538 IGEQAIGSDRDRVVRISTLPETGARG 563

RESULT 11

A48562

Db	145	EPSAQWSTAADMASGKSVSEW-----EAFPSFHTSVNMWSTSETQCKILFKQSLGPLLN	198
Qy	87	PYLAHLARMYNGHAGMEVQIVLAGNAFTAGKIIFAAIPGGF-PYENLSPSQITMCPHVI	145
Db	199	PYLEHLKSLKYAVWSSGSIIEVRFSISGSGVFGKKLAAIIVPPGVDPVQSTMLQY---PHVL	255
Qy	146	IDVRQLPEPFLPMPDINWNFFHYNQGNQDKPLRLVAMLYTPL-----RANNSGDDVFTVSC	200
Db	256	FDAQVEPFIPTPDLRSTLYHWMSDTD-TTSLVIVVYNDLINPYANDSNSSGCIIVTE-	313
Qy	201	RVLTKSPDPEFTFLVPPP-----	218
Db	314	---TKPGDPDFKFLKPPGSSVLTHGSIPIPSDLIPKSSSLWIGNRYWTDITDFVIRPFVQA	370
Qy	219	-----TVESKTKQFALPIKISEMTNSRPPVPVDMVY-----TARNENQV	258
Db	371	NRHDFNQETAGWSTPRFRITITISEKNGSKLIGVATDYIIPGIPDGPWDTTIADKLI	430
Qy	259	VQPQNGRVTLJDGELLGT-----TPLLAVNICFKPGKEVIAGKGVRSVRMDMEITNT----	309
Db	431	PAGDYSITTCGEINDIKTAQAYDTAAVVKNTTFRGMVIC--GSLQRAWGDKKISNTAFIT	488
Qy	310	-----DGTPIDPT	317
Db	489	TAIRDGNEIKPS	500
RESULT 13			
VCWVFF			
coat protein - feline calicivirus (strain CFI/68 FIV)			
N:Alternate names: capsid protein			
C:Species: feline calicivirus			
C:Date: 30-Jun-1992 #sequence revision 30-Jun-1992 #text_change 20-Aug-1999			
C:Accession: A40507; B40507; T09246			
R:Neill, J.D.; Reardon, I.M.; Heinrikson, R.L.			
J. Virol. 65, 5440-5447, 1991			
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline			
A:Reference number: A40507; MUID:91374597; PMID:1716692			
A:Accession: A40507			
A:Molecule type: genomic RNA			
A:Residues: 1-668 <NEI>			
A:Cross-references: GB:M32819; NID:g323874; PIDN:AAA42925.1; PID:g323875			
A:Accession: B40507			
A:Molecule type: Protein			
A:Residues: 373-379; 403-419; 481-489; 560-566 <NE2>			
submitted to the EMBL Data Library, April 1998			
A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68			
R:Neill, J.D.			
A:Reference number: Z16626			
A:Accession: T09246			
A:Status: preliminary; translated from GB/EMBL/DBJ			
A:Molecule type: genomic RNA			
A:Residues: 1-668 <NE3>			
A:Cross-references: EMBL:U13992; NID:g3056875; PIDN:AAC13993.1; PID:g537256			
A:Experimental source: strain CFI/68 FIV			
C:Superfamily: feline calicivirus coat protein			
C:Keywords: capsid protein; coat protein; glycoprotein			
F:177,301,439,459,615/Banding site: carbohydrate (Asn) (covalent) #status F			
Query Match 8.8%; Score 254; DB 1; Length 668;			
Best Local Similarity 36.7%; Pred. No. 1.9e-11;			
Matches 58; Conservative 26; Mismatches 60; Indels 14; Gaps 5			
Qy	67	SPRNSPGIILDLGPDNLNLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKIIFAAIAPP	126
Db	179	STSEFQCKILFKQSLGPLNLNLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKIIFAAIAPP	238
Qy	127	GF-PYENLSPSQITMCPHVIDVRQLEPFLPMPDINWNFFHYNQGNQDKPLRLVAMLYTYP	185
Db	239	GIDPVQSTMLQY---PHVLFDAQVEPFIPTPDLRSTLYHWMSDTD-TTSLVIVVYND	294
Qy	186	L-----RANNSGDDVFTVSCVRLTKPSPDFEFTFLVPP	218

Db 295 LINPYANDSNSSGCVITVE-----TKPGDFKFHLLKPP 328

RESULT 14
JQ2354
capsid protein - feline calicivirus (strain NADC)
C/Species: feline calicivirus
C/Date: 28-Aug-1995 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C/Accession: JQ2354
R/Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A/Title: Analysis of feline calicivirus capsid protein genes: identification of variable
A/Reference number: JQ2354; MUID:94065683; PMID:7504075
A/Accession: JQ2354
A/Molecule type: mRNA
A/Residues: 1-668 <SEA>
A/Cross-references: GB:I09718; NID:g305104; PID:AAA16485.1; PID:g305105
C/Superfamily: feline calicivirus coat protein
C/Keywords: capsid protein; coat protein

Query Match 8.6%; Score 248.5; DB 2; Length 668;
Best Local Similarity 32.8%; Pred. No. 5.1e-11;
Matches 65; Conservative 32; Mismatches 78; Indels 23; Gaps 7;

Qy 30 EPVAGASLPTVVGQNIIDPWIRNNFVQAPAGEFTV---SPRNSPGEILLDLLEGLDPLN 86
Db 145 EPSAQMSTAADMATGKSDSEW-----EAFPSHTSVNMSTSETQGKILFKQSLGSLLN 198
Qy 87 PYLAHLARMYNGHAGGMEVOIVLAGNAFTAGKIIFRAIPGF-PYENLSFSQITMCPHVI 145
Db 199 FYLEHLSKLVAMSGSVVEFRFSISGSGVGGKLAALIVVPFGVDFVQSTSLQY---PHVL 255
Qy 146 IDVRQLPEPFLPPDIDWNNEFFHYNQGNDPKLRVLVAMLYTPL-----RANNSGDDVFVTSV 200
Db 256 FDAQVDVFVFSIPDLSTLYHLMPDTD-TTSLVIMVYNDLINPYANDSNSSGCVITVE- 313
Qy 201 RVLTKPSPDFEFTFLVPP 218
Db 314 ---TKPGDFKFHLLKPP 328

```

RESULT 15
JQ2356
capsid protein - feline calicivirus (strain KCD)
C:Species: feline calicivirus
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: JQ2356
R:Seal, B.S.; Riepath, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable
A:Reference number: JQ2354; MUID:94065683; PMID:7504075
A:Accession: JQ2356
A:Molecule type: mRNA
A:Residues: 1-668 <SEA>
A:Cross-references: GB:L09719; NID:G305107; PIDN:AAAL6487.1; PID:G305108
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein

Query Match      8.6%; Score 248.5; DB 2; Length 668;
Best Local Similarity 24.5%; Pred.No.5.1e-11;
Matches 112; Conservative 66; Mismatches 171; Indels 109; Gaps 22;

Qy      12  SDGAAGLVPRINNEV---MPLPVPAGASLATPFVVGQNIIDPWIRNNFVQAPAGEFTV--- 66
Db      125  ADGGSIVTPGQGTWVGGVIAEPNAQMSADMATGKSDVSEW-----EAFSFTSVNW 178

Qy      67  SPRNSPGIEILLDELGLPDLNPLYLAHARMYTNHAGGMEVQIVLAGNAFTAGKIIFAAI 136
Db      179  STSETQKILPKQSLGLPDLNPLYLAHARMYTNHAGGMEVQIVLAGNAFTAGKIIFAAI 136

Qy      127  GP-PYENLSPSQITMCPHVIVDVRQLEPFLPMPDIWNPNFFHYNQGNPKLRILVAMLYTP 185
Db      239  GVDPEVOSTSMLOV---PHVLFDARQVEPVFISIDLRSLTLYHLMSTDT-TTSLVMWAYND 294

```

THIS PAGE BLANK (USPTO)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	1563	54.0	539	1	COAT LORDV	P54635 lordsdale v
2	966	33.4	546	1	COAT SOUV3	O04542 southampton
3	311	10.7	2344	1	POLN RHUV	P27410 rabbit hemo
4	271.5	9.4	703	1	COAT SNSV4	P36285 san miguel
5	262.5	9.1	671	1	COAT SCVF9	P27406 feline cali
6	262	9.1	702	1	COAT FSNV4	P36284 san miguel
7	260	9.0	668	1	COAT FCVF4	P27405 feline cali
8	254	8.8	668	1	COAT FCVC6	P27404 feline cali
9	247.5	8.6	2208	1	POLN MNCV	Q69014 manchester
10	126.5	4.4	2194	1	POLG HE701	P32537 h genome po
11	125.5	4.3	2193	1	POLG CX16T	Q9Q131 c genome po
12	124.5	4.3	2193	1	POLG CX16G	Q65900 c genome po
13	121.5	4.2	2184	1	POLG EC01F	O91734 e genome po
14	120	4.1	2193	1	POLG HE71M	Q66479 h genome po
15	118.5	4.1	2333	1	POLG FMDV1	P03306 f genome po
16	118	4.1	2183	1	POLG CXB4E	O86887 c genome po
17	118	4.1	2193	1	POLG HE71B	Q66478 h genome po
18	117.5	4.1	2205	1	POLG POL2W	P23069 p genome po
19	117	4.0	2175	1	POLG BOVEV	P12915 b genome po
20	115.5	4.0	2206	1	POLG POL1M	P03299 p genome po
21	115.5	4.0	2208	1	POLH POL1M	P03300 p genome po
22	113.5	3.9	2206	1	POLG CXA21	P22055 c genome po
23	112.5	3.9	2203	1	POLG TMEVG	P08545 c genome po
24	112	3.9	2185	1	POLG SVDHV	P16604 c genome po
25	112	3.9	2301	1	POLG TMEVD	P13899 t genome po
26	112	3.9	2332	1	POLG FMDVA	P03308 f genome po
27	111	3.8	2183	1	POLG CXB4J	P08292 c genome po
28	110.5	3.8	862	1	POLG EC16H	Q66790 echovirus 1
29	110	3.8	2209	1	POLG POL1S	P03301 p genome po
30	109	3.8	2185	1	POLG SVDVU	P13900 s genome po
31	109	3.8	2290	1	POLG ENCV	P03304 e genome po
32	108.5	3.7	2207	1	POLG POL2L	P06210 p genome po
33	106.5	3.7	2194	1	POLG EC30B	Q9wn78 e genome po

```

QY 181 MLYTPLRANNSGDDVFTVSCRVLTKPSPDFEFLVPPVTSKTKQFALPILKISEMTNS 240
DB 181 MLYTPLRANNSGDDVFTVSCRVLTKPSPDFEFLVPPVTSKTKQFALPILKISEMTNS 240
QY 241 RFPVPVDMYMTARNENQVVPQNGRVTLDGELLGTTPLLAIVNICKFKGVIKNGDVRSY 300
DB 241 RFPVPVDMYMTARNENQVVPQNGRVTLDGELLGTTPLLAIVNICKFKGVIKNGDVRSY 300
QY 301 RMDMEITNDGTPIEDTTPGIPGSPDFQGLIFGVASORNNKNEQNAPTRAHEAINTGG 360
DB 298 DYTMLASQWNSYDPTBPAIPAGTPGDFVGKIQGLLTQTR--ADGSTRAHKATVSTGS 355
QY 361 DHLCPQISSEIYLTSNLRCTNPOPLQSGGLRGTLIRSDNGCHDMVGSPITPTWP 420
DB 356 VHFPTKLGVSQVFTDNTNDFQAGNTKFTFVG-----VIQDGDHQQNE-----P 399
QY 421 QQWRRCSRGNSCCSGHRYP--VPVVMNRVTWIVLSHKSGFSTSKLPLQNL-----R 472
DB 400 QWSLPNYSGRCHNVHLAPAVAPTFFGQLLFFRSTMPGCS-----GYPNMLDCLLPQE 455
QY 473 W-----PLIRFNPDTGRVLFEARLHKQGITVAHTGDNPIVMPNGYPR 517
DB 456 WVLHFYQEAAPAQSDVALLRFVNPDTGRVLFECKLHKSGYITVAHTGPDYDLVLPNGYPR 515
QY 518 FEAWNQFYSLAPVGTGKGRRR 539
DB 516 FDSVWNQFYTLAPMGNGTGRRR 537

RESULT 2
COAT_SOUV3 STANDARD; PRT; 546 AA.
ID COAT_SOUV3
AC Q04542;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Coat protein (Capsid protein).
OS Southampton virus (serotype 3).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=37129;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93142023; PubMed=8380940;
RA Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;
RT "Sequence and genome organization of a human small round-structured (Norwalk-like) virus.";
RL Science 259:516-519(1993).
CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L07418; AAA92984.1; -.
DR PIR; B37491; B37491.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 546 AA; 58774 MW; C82B2A85AD4B05EA CRC64;
Query Match 33.4%; Score 966; DB 1; Length 546;
Best Local Similarity 38.5%; Pred. No. 9.8e-66;

```

```

Matches 225; Conservative 87; Mismatches 189; Indels 84; Gaps 17;
QY 1 MKMASNDAAPSDGAAG--LVPEINN-EVMPLEPVAGASLATPVVGQNIIDPIWRNWF 56
DB 1 MKMASNDAAPSDGAAG--LVPEINN-EVMPLEPVAGASLATPVVGQNIIDPIWRNWF 56
QY 57 VQAPAGEFTVSPRNSPGEIILDLGLPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAPTA 116
DB 61 VQSPQGEFTISPNNTPGDILFDLQLGPHLNPFLSHLSQMYNGWGNMVRVILLAGNAPSA 120
QY 117 GKILFAALPPGFPPENLSPQITMCPHVIVDROLEPFLBMPDIWNFFHYVNGNDPKL 176
DB 121 GKIIIVCCVPPGFTSSSLTIAQTILFPHIADVTRLEPIEMPLEDRVNVLYHTND-NQDTM 179
QY 177 RLVAMLYTLPLR--ANNSGDDVFTVSCRVLTKPSPDFEFLVPPVTSKTKQFALPILKI 234
DB 180 RLVCMLYTLPLRTGGSGNSDSFVAGRVLTAPSDFSFLVLPPTIEQKTRAFVNPNIPL 239
QY 235 SEMTNSRFPVDMYMTARNENQVVPQNGRVTLDGELLGTTPLLAIVNICKFKGVIKNG 294
DB 240 QTLNSRFPSPSLIQGMILSPDASQVVQFQNGRCLIDGQLLGTTPATSGQLFRVRGKI--N 296
QY 295 GDVRSYRMDMEITNDGTPIEDTTPGIPGSPDFQGLIFGVASORNNKNEQNAPTRAHEA 354
DB 297 QGART-----LNLTEVDGKPF--MAFDSPAPVGFDPFGKCDWMHRISKTPNTTSGDPMRSV 351
QY 355 IINTGGDHLCPQISS---SEIYLTSPNLRCTNPOPLQSGGLRGTLIRSDNGCHDMVG 411
DB 352 SVQTNVQGVFVHLGSIQFDEVF-----NHPTGDYIG 382
QY 412 T-----SPTTP--TWPOQWRRCSRGNSCCSGHRYPVPVVMNRVTWIVLSHKSGFSTSTR 464
DB 383 TIEMISQPTPPTDINLWEIPDYGSSLSQAANLAP-PVFPFGFGEALVYFVSAPPGPN 441
QY 465 K-----LPO-----LNLRW-----LIRFNPDTGRVLFEARLHKQGITVA-- 501
DB 442 RSAPNDVPCLLPQEIYTHFVSEQATMGDAALLHYVDPDTRNLNGLFEFLPGGYLTCPVN 501
QY 502 HTGDNPIVMPNGYPRFEAWNQFYSLAPVGTGK-----GRRRV 540
DB 502 GVGAGPQQLPLNGVFLFVSWSRFFQLKPVGTASTARGRLGVRI 546

RESULT 3
POLN_RHDV STANDARD; PRT; 2344 AA.
ID POLN_RHDV
AC P27410;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Non-structural polyprotein [Contains: RNA-directed RNA polymerase
DE (EC 2.7.7.48); Thiol protease P3C (EC 3.4.22.-); Helicase (2C like
DE protein); Coat protein].
OS Rabbit hemorrhagic disease virus (RHDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Lagovirus.
OX NCBI_TaxID=11976;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91361557; PubMed=1840711;
RA Meyers G., Wirblich C., Thiel H.-J.;
RT "Rabbit hemorrhagic disease virus -- molecular cloning and nucleotide
RT sequencing of a calicivirus genome.";
RL Virology 184:664-676(1991).
CC -1- FUNCTION: THE P3C CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE
CC FOR CLEAVAGE AT CERTAIN Q/G SITES IN THE POLYPROTEIN.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```


QY 63 EFTV-----SPRSPGEILLDLGPDLPNLYLAHARMYNGHAGMEVOIVLAGN 112
 Db 196 EMTTFSYHTANVMSSTEAQKILFSRALSPYLRHISLSYTWSGGIDVRFVSGS 255
 QY 113 AFTAGKIIFAAIPPGF-PYENLSPSOITWCPHVIIIDVROLEPPDLLPMPDIWNNFFHYNOG 171
 Db 256 GVFGKLAALIVPPGIEVE--SPTML-QYPHVLFDAQTQEPVIFIPDIRKTLYHSMDD 312
 QY 172 NPKRLVLAWLTYTPL-RANNSGDDVFTVSCRVLTKPSDPFEFTFLVPPTVESKTKQFALP 230
 Db 313 TD--TTRLVIMVYNELINPEQSEPKSSCSITVETRPSSDFTSLKPP--GSLKHGSI 369
 QY 231 ILKISEWTSRFPVVDVMTARN-----ENQVQPOQ---NGRVTLDGELLG-TTPL- 278
 Db 370 ----SOL-----IPRNSRHMGNRWMSITDGFVQVFRFQSNRHRPFDFTTGMSTPY 419
 QY 279 --LAVNICKFKGEVIAKNGDVRSYRMDMEITNTDGTPIDPTEDTPG-PIGSPD 328
 Db 420 IPIEVL-----EKLDGGQY-----FKVTDEKSLV-----PGLPDGWF 455

RESULT 7

COAT_FCVF4
 ID COAT_FCVF4 STANDARD; PRT; 668 AA.
 AC P27405;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Coat protein (Capsid protein).
 OS Feline calicivirus (strain Japanese F4) (FCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OX NCBI_TaxID=11980;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91306470; PubMed=1853578;
 RA Tomya Y., Taniguchi Y., Takahashi E., Utagawa E., Takeda N.,
 Miyamura K., Yamazaki S., Mikami T.;
 RT "Sequence analysis of the 3'-end of feline calicivirus genome.";
 RL Virology 183:810-814(1991).
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D90357; BAAL4371.1; -.
 DR PIR; B40481; VCMWFC.
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 668 AA; 73589 MW; 85BDBC85804B503 CRC64;

Query Match 9.0%; Score 260; DB 1; Length 668;

Best Local Similarity 25.0%; Pred. No. 4.9e-12;

Matches 93; Conservative 46; Mismatches 133; Indels 100; Gaps 12;

QY 30 EPVAGSLATPVVGOQNIIDPWIRNFFVOAPAGEFTV---SPRNSGETLLDLGLGPDLL 86
 Db 145 EFSAQNSTAADNASKGVDSW-----EAFSFTSVNWSVSTQGKILFKQSLGPDLL 198

QY 87 PYLAHARMYNGHAGMEVOIVLAGNAFTAGKIIFAAIPPGF-PYENLSPSOITWCPHY 145

Db 199 PYLEHLSKLYVAMSGSIEVRSISGSGVFGGKLAIVVPPGVDVQVSTSMLOY----PHVL 255
 QY 146 IDVROLEPFLPMPDIWNNFFHYNOGNDPKLRLVLAWLTYTPL-----RANNSGDDVFTVSC 200
 Db 256 FDAQVEPVIFIPDLRSLTYHVMSDTD--TTSLVIMVYNLDLINPYANDSNSGCCIVTVE- 313
 QY 201 RVLTKPSDDPFEFTFLVPP----- 218
 Db 314 ---TRPGDPDKFHLKPKGSLVTHGSIIPSDLIIPKSSSLMIGNRYWYTDITDFVIRPFVFOA 370
 QY 219 -----TVESKTKQFALPILKISEMNTSNRFPVDMV-----TARNENOV 258
 Db 371 NRHFDNFOETAGSTPRFRPITITISEKNGSKLGIGVATDYIIPGLPDGWPDTTIADKLI 430
 QY 259 VQOPNGRVTLDGELLGT-----TPLLAVNICKFKGEVIAKNGDVRSYRMDMEITNT 309
 Db 431 PAGDYSIITTEGENDIKTAQAYDTAAVKNVNTNFRGMWYC--GSLQRAWGCDKISNTAFIT 488
 QY 310 ----DGTPIDPT 317
 Db 489 TAIRDGNEIKPS 500

RESULT 8

COAT_FCVF6
 ID COAT_FCVF6 STANDARD; PRT; 668 AA.
 AC P27404;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Coat protein (Capsid protein).
 OS Feline calicivirus (strain CFI/68 FIV) (FCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OX NCBI_TaxID=11979;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=91374597; PubMed=1716692;
 RA Neill J.D., Reardon I.M., Heinrichson R.L.;
 RT "Nucleotide sequence and expression of the capsid protein gene of
 RT feline calicivirus";
 RL J. Virol. 65:5440-5447(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Neill J.D.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- PTM: The N-terminus is blocked.
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M32819; AAA42925.1; -.
 DR EMBL; U13992; AAC13993.1; -.
 DR PIR; A40507; VCMWFF.
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 668 AA; 73550 MW; 9E52312108D4D41 CRC64;

Query Match 8.8%; Score 254; DB 1; Length 668;
Best Local Similarity 36.7%; Pred. No. 1.4e-11;
Matches 58; Conservative 26; Mismatches 60; Indels 14; Gaps 5;
QY 67 SPNSGGEILLDLLELPGDLPNLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKIIFAAIIPP 126
Db 179 STSETGKGLFKQSLGGLPPLNLYLTHLAKLYVAVSGVDVRFSGSGVFGKLAIVVPP 238
QY 127 GP-PYENLSPSQITMCPHVIDVRQLEPFLPMPDIWNNFFHYNQNDPKLILVAMLYTP 185
Db 239 GIDPVQSTMLQY---PHVLFDAQVERPFIIPDLRSTLYHLMSDTP-TTSLVIMVND 294
QY 186 L-----RANSGDDVFTVSCRVLTKPSPDFEFTFLVPP 218
Db 295 LINPYANDSSGCVITVE-----TKGPDFKFHLKPP 328

RESULT 9
POLN_MANCV STANDARD; PRT; 2208 AA.
ID POLN_MANCV STANDARD; PRT; 2208 AA.
AC Q69014;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: RNA-directed RNA polymerase
(EC 2.7.7.48); Thiol protease 3C (EC 3.4.22.-); Helicase (2C like
protein); Coat protein].
DE Manchester virus (Human enteric calicivirus).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Sapovirus.
OX NCBI_TaxID=82659;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95390791; PubMed=7661689;
RA Liu B.L., Clarke I.N., Caul E.O., Lambden P.R.;
RT "Human enteric caliciviruses have a unique genome structure and are
distinct from the Norwalk-like viruses.";
RL Arch. Virol. 140:1345-1356(1995)
CC -!- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).
CC -!- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR
THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY
SIMILARITY).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
{RNA}(N).
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

EMBL; X86560; CAA60262.1; --
DR MEROPS; C24.0UP; --
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR000317; Peptidase_C24.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR Pfam; PF03510; Peptidase_C24; 1.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00916; 2CENDOPTASE.
DR PRINTS; PR00918; CALICIVIRUSN.
KW polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;

KW Thiol protease; Helicase; ATP-binding; Coat protein.
FT CHAIN ? ? HELICASE (P2C-LIKE).
FT CHAIN ? ? PROTEASE (P3C).
FT CHAIN ? ? RNA-DIRECTED RNA POLYMERASE.
FT CHAIN ? ? COAT PROTEIN.
FT NP_BIND 408 415 ATP (POTENTIAL).
FT ACT_SITE 1097 1097 PROTEASE (BY SIMILARITY).
FT ACT_SITE 1112 1112 PROTEASE (BY SIMILARITY).
SQ SEQUENCE 2208 AA; 242736 MW; 3E299D5BA20E45DC CRC64;
Query Match 8.6%; Score 247.5; DB 1; Length 2208;
Best Local Similarity 32.4%; Pred. No. 2.2e-10;
Matches 73; Conservative 31; Mismatches 104; Indels 17; Gaps 6;
QY 2 KMASNDAAAPSDGAAGLVPEINNEVM--PLEPVAGA---SLATPVVGQNIIDPWIRNMF 56
Db 1658 KQSNPMVVDPPGTG--PTTSHVVVANPEQPGNAAQBLEAVATGAIQSNVPEAIRNCF 1715
QY 57 VQAPAGEFTVSPRNSPGEILLDLLELPGDLPNLYLAHLARMYNGHAGMEVQIVLAGNAFTA 116
Db 1716 --AVPRTFAWDRMPTGTGFLGSLSLHPNINPYTSHLSGMWAGWGSFEVRLISGSGVFA 1773
QY 117 GKILPAALPPGEPYENLSPSOIT---MCPHVIDVRQLEPFLPMPDIWNNFFHYNQND 173
Db 1774 GRITASVIPP-----VDPSSIRDPGVLPHAFVDAIRITEPVSFMIIPDVAVDYHRMDGAE 1828
QY 174 PKLRLVAMLYTPLRANSGDDVFTVSCRVLTKPSPDFEFTFLVPP 218
Db 1829 FTCSLGFVWYQDLPNLPFTTAVSICWSVETKPGGDFDFCLLRPP 1873

RESULT 10
POLG_HE701 STANDARD; PRT; 2194 AA.
ID POLG_HE701 STANDARD; PRT; 2194 AA.
AC P32537;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (PIA); Coat protein VP2
(PIB); Coat protein VP3 (PIC); Coat protein VP1 (PID); Core protein
P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-
linked protein VP6 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)
(P3C); RNA-directed RNA polymerase P3D (strain J670/71) (EV 70).
OS Human enterovirus 70 (strain J670/71) (EV 70).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=31915;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91037960; PubMed=2172447;
RA Ryan M.D., Jenkins O., Hughes P.J., Brown A., Knowles N.J., Booth D.,
RA Minor P.D., Almond J.W.;
RT "The complete nucleotide sequence of enterovirus type 70:
relationships with other members of the picornaviridae.";
RL J. Gen. Virol. 71:2291-2299(1990).
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the
poliovirus polypeptide. In other picornavirus reactions Glu may be
substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
{RNA}(N).
CC each of which is composed of one copy each of proteins VP1, VP2,
VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
CC or send an email to license@isb-sib.ch).

```
DR EMBL; D00820; BAAL1891.1; -.
DR PIR; A36253; GNNVE7.
DR HSSP; P03300; IPOV.
DR MEROPS; C03.UFA; -.
DR MEROPS; C03.UFB; -.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR000199; Pept_3C_picorn.
DR InterPro; IPR000081; Peptidase_C3.
DR InterPro; IPR003138; Pico_P2B.
DR InterPro; IPR002527; Pico_P2B.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR00605; RNA helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF0548; Cys-protease_3C; 1.
DR Pfam; PF02226; Pico_P1A; 1.
DR Pfam; PF00947; Pico_P2A; 1.
DR Pfam; PF01552; Pico_P2B; 1.
DR Pfam; PF00073; rhv; 3.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICIVIRUS.
DR ProDom; PD001125; Cys_protease_3C; 1.
DR ProDom; PD001306; Pico_P2A; 1.
DR ProDom; PD001274; Pico_P2B; 1.
DR SMART; SM00382; AAA; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
FT CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 319 COAT PROTEIN VP2.
FT CHAIN 320 561 COAT PROTEIN VP3.
FT CHAIN 562 871 COAT PROTEIN VP1.
FT CHAIN 872 1014 CORE PROTEIN P2A.
FT CHAIN 1015 1113 CORE PROTEIN P2B.
FT CHAIN 1114 1443 CORE PROTEIN P2C.
FT CHAIN 1444 1532 CORE PROTEIN P3A.
FT CHAIN 1533 1554 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1555 1737 PICORNAIN 3C.
FT CHAIN 1738 2194 RNA-DIRECTED RNA POLYMERASE P3D.
FT LIPID 2 2 N-myristoyl glycine (in host) (By
FT ACT SITE 1701 1701 PROTEASE (POTENTIAL).
FT ACT SITE 1715 1715 PROTEASE (POTENTIAL).
SQ SEQUENCE 2194 AA; 244590 MW; 15DBAE96EE06673C CRC64;
```

Query Match 4.4%; Score 126.5; DB 1; Length 2194;
Best Local Similarity 23.0%; Pred. No. 0.34;
Matches 65; Conservative 38; Mismatches 94; Indels 85; Gaps 13;

```
QY 21 EINNEWPLEPVAGASLAPVVGQQNIIDPWIRNNFVQAPAGEFTVSPRNSPGEILLDL 80
DB 373 EINN-----VNDASGVRLRVISAQSDMDQLFN-----IPLDIQ 408
QY 81 L-GDPLNPYLAHLMYNGHAGMEVQIIVLAGNAFTAGKIIIFAAIPPGPPYENLSPS--- 136
DB 409 LSGPLRNTLLGNISRYVTHWSGSLEMTFMCSPFTTKGLIICYTPPG-----GSSPTDRM 464
QY 137 QITMCHVIVDVRQEPFLPMPDIWNFFHYNQGNPKRLVRLVAMLYPLRANNSGDDVF 196
DB 465 QAMLATHVVDVDFGLQSSITIIIP--WISGSHYR-----MFTDAKAINA--NVG 509
QY 197 TVSCRVLTSPDFEFTFLVPT-----VSKTKQFALPILKISEMNTSPFPVPV 246
DB 510 YVTC-----FMQTNLVAFGAADCQYIVGMVAACKDFNLRMRDSP----- 550
```

```
QY 247 DVMYTARNENQVVPQNGRV-----TLDGEL---LGTTPLL 279
DB 551 DIGQSAILPEQAATTQIGIEIVKTVANTVESEIKAEGLVIPS 592

RESULT 11
POLG_CX16T STANDARD; PRT; 2193 AA.
AC Q9QF31;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VPG (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].
DE Cxsackievirus A16 (strain Tainan/5079/98).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.
OC NCBI_TaxID=231417;
RX MEDLINE=21427116; PubMed=11536241;
RA Yan J.-J., Su I.-J., Chen P.-F., Liu C.-C., Yu C.-K., Wang J.-R.;
RT "Complete genome analysis of enterovirus 71 isolated from an outbreak in Taiwan and rapid identification of enterovirus 71 and coxsackievirus A16 by RT-PCR.";
RL J. Med. Virol. 65:331-339(2001).
CC -!- FUNCTION: It is thought that the P2C protein attaches to vesicular membranes and is associated with viral RNA synthesis.
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-[Gly bond in the poliovirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA} (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
```



```
DR Pfam; PF00910; RNA helicase; 1.
DR PRINTS; PR00918; CALICIVIRUSNS.
DR ProDom; PD001125; Cys protease_3C; 1.
DR ProDom; PD001306; Pico_P2A; 1.
DR ProDom; PD001274; Pico_P2B; 1.
DR SMART; SM00382; AAA; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
KW Lipoprotein.
FT CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 323 COAT PROTEIN VP2.
FT CHAIN 324 565 COAT PROTEIN VP3.
FT CHAIN 566 862 COAT PROTEIN VP1.
FT CHAIN 863 1012 CORE PROTEIN P2A.
FT CHAIN 1013 1111 CORE PROTEIN P2B.
FT CHAIN 1112 1440 CORE PROTEIN P2C.
FT CHAIN 1441 1526 CORE PROTEIN P3A.
FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VP6.
FT CHAIN 1549 1731 PICORNAIN 3C.
FT CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE.
FT LIPID 2 2 N-myristoyl glycine (in host) (By
FT ACT_SITE 1695 1695 PROTEASE (POTENTIAL).
FT ACT_SITE 1709 1709 PROTEASE (POTENTIAL).
FT SEQUENCE 2193 AA; 243182 MW; 927839DB58F61E7F CRC64;

Query Match
Best Local Similarity 4.3%; Score 125.5; DB 1; Length 2193;
Matches 88; Conservative 49; Mismatches 169; Indels 93; Gaps 17;

QY 1 MKMASNDAAAPSSDG-AAGLVP-----EINN--EVMPLPVPAGA-SLATPVVQ 44
Db 329 LKPGTNQFLTTDDGVSAPIPLGPHPTPTTHIPGEVHNLEICREVTEILEVNNLKT----- 383

QY 45 QNIIDPWIRNF---VQAPAGFTVSPRNSPGEILLDLDELGPDLNPLYLAHARMYNGHAG 101
Db 384 -NETTFMQRLCFVSVQSKTGELCAAFRADPGRD-----GPMQSTILGQLCRYTQWSG 436

QY 102 GNEVQIVLAGNAFTAGKIIFAAIPGCFYENISPSQIT--MCPHVIDVRQLEPFLPMP 159
Db 437 SLEVTFMFAGSWATQGMIAIATYPPG---GNVPADRITAMLGHVTLWDFGLQSSVTLVVP 493

QY 160 DIWNFFHHYNOGNDPKRLRLVAMLYTPLRANNSGDDVFTVSCRVLTKPSDFEFTFLVP-- 217
Db 494 -WISNTHYR-----AHARAGYFDVYTTGIITW-----YQTNVYVPIG 530

QY 218 -PT-----VESKTKQFALPILKISEMTNSRPV-----PVDVMYATARNQV-----V 259
Db 531 APTTAYIVALAAQDNFTMKLCKDTEIDQTANIQGDPIADIMIDQTVNNQVNRSLTALQV 590

QY 260 QPQNGRVTLDGELLGTTPLLAVNICFKGEVIAKNGDVRSYRMDMEITWTGDTGTPID- 315
Db 591 LPTAADTEASSHRLGTVGVVPLQAQASTGASSNASDKNLIETRCVLNHHSTQETAIQNFPS 650

QY 316 -----PTEDTPGPGSPDFQGLFGVASQRNKN 344
Db 651 RAGLVSIITMPTTGITQNTDGYVNDWIDLNGYAQLRKCE 689

RESULT 12
POLG_CX16G
ID POLG_CX16G STANDARD; PRT; 2193 AA.
AC Q65900;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2
DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein
DE P2A; Core protein P2B; Core protein P3A; Core protein P3B; Genome-
DE linked protein VP6 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].
OS Coxsackievirus A16 (strain G-10).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
```

```
OC Enterovirus.
RN [1] NCBI_TaxID=69159;
RP SEQUENCE FROM N.A.
RA MEDLINE=94303216; PubMed=8030260;
RX Poyry T., Hyypia T., Horsnell C., Kinnunen L., Hovi T., Stanway G.;
RT "Molecular analysis of coxsackievirus A16 reveals a new genetic group
of enteroviruses.";
RL Virology 202;982-987(1994).
CC -!- FUNCTION: It is thought that the P2C protein attaches to vesicular
membranes and is associated with viral RNA synthesis.
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the
poliovirus polyprotein. In other picornavirus reactions Glu may be
substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
[RNA] (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
each of which is composed of one copy each of proteins VP1, VP2,
VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U05876; AAA50478.1; -.
DR HSSP; P03300; IPOV.
DR MEROPS; C03.022; -.
DR MEROPS; C03.UFA; -.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR004004; Calici pol hel.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR000199; Pept_3C_picorn.
DR InterPro; IPR000881; Peptidase_C3.
DR InterPro; IPR003138; Pico_P1A.
DR InterPro; IPR002527; Pico_P2B.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR006005; RNA helicase.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR001205; RNA pol_P3D.
DR InterPro; IPR007094; RNA pol_Psvir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00548; Cys-protease-3C; 1.
DR Pfam; PF02226; Pico_P1A; 1.
DR Pfam; PF00947; Pico_P2A; 1.
DR Pfam; PF01552; Pico_P2B; 1.
DR Pfam; PF00073; rhv; 3.
DR Pfam; PF00680; RNA dep RNA_pol; 1.
DR Pfam; PF00910; RNA helicase; 1.
DR PRINTS; PR00918; CALICIVIRUSNS.
DR ProDom; PD001125; Cys protease_3C; 1.
DR ProDom; PD001306; Pico_P2A; 1.
DR ProDom; PD001274; Pico_P2B; 1.
DR SMART; SM00382; AAA; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
KW Lipoprotein.
FT CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 323 COAT PROTEIN VP2.
FT CHAIN 324 565 COAT PROTEIN VP1.
FT CHAIN 566 862 COAT PROTEIN VP3.
FT CHAIN 863 1012 CORE PROTEIN P2A.
FT CHAIN 1013 1111 CORE PROTEIN P2B.
FT CHAIN 1112 1440 CORE PROTEIN P2C.
FT CHAIN 1441 1526 CORE PROTEIN P3A.
FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VP6.
```


FT CHAIN 1549 1731 PICORNAIN 3C.
FT CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE.
FT LIPID 2 N-myristoyl glycine (in host) (By
FT similarity).
FT ACT SITE 1695 1695 PROTEASE (POTENTIAL).
FT ACT_SITE 1709 1709 PROTEASE (POTENTIAL).
SQ SEQUENCE 2193 AA; 243209 MW; 04B3BCE572A76E38 CRC64;
Query Match
Best Local Similarity 41.3%; Score 124.5; DB 1; Length 2193;
Matches 86; Conservative 47; Mismatches 179; Indels 81; Gaps 15;
QY 1 MKVANSNDAPSDG--AAGLVP-----EINNEVMPLEPVAGASLATPVG-QQNIIDP 50
Db 329 LKPGTNQFLTDDGVSAPILPGFHPPTPAIHGPEVRNLLIEICRVETILEVNNLSNETTP 388
QY 51 WIRNNF---VQAPAGEFTVSPNRSCEILLDLLELGLDLPNLYLAHARMYNGHAGGMEVOI 107
Db 389 MORLCFPVSVQSKTGELCAVFRADPG-----RNGFWQSTILGQLCRYTQWSGSLEVT 442
QY 108 VLAGNAFTAGKIFRAIPP--GPPYENLSPSQTMCPHVIIDVRQLEPFLPMPDINWNNF 165
Db 443 MFAGSPMATGKMLIATVPGGVGPADRLTAMLTG---HVIDFGLQSSVTLVIP--WISN 497
QY 166 FHYNQNDPKLRILVAMLYTPLRANNSGDVFTVSCRVLTKPSDFEFTFLVP---PT--- 219
Db 498 THYR-----AHAKGYFDYTTTGTITW-----YQTNVYWPFGAPTAY 536
QY 220 ---VSKTKQFALPILKISEMINSFPVP-----VDVMTARNENQV-----VQPNQR 265
Db 537 IVALAAQDNFTMKLCKDTEIEQSANIQGGDIADMDQAVTSRVRGALTSIQVEPTAAN 596
QY 266 VTLDGELLGTTPLLVNICKRGEVIAKNQGVRSYRMDMEITNDCGTPID----- 315
Db 597 TNASEHRLGTGLVPAQAAETGASSNAQDENLIEFVCLNHHSTQETTIGNFFSRAGLV 556
QY 316 ----PTEDTPGPIGSPDFGILFVGASQNKNE 344
Db 657 IITMPTGTQNTDGYVNWIDLMGVAMQRKCE 689
RESULT 13
POLG EC01F STANDARD; PRT; 2184 AA.
AC DT 15-JUL-1999 (Rel. 38, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein (Contains: Coat protein VP1 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Picornain 2A (EC 3.4.22.29) (P2A); Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)).
OS Echovirus 1 (strain Farouk / ATCC VR-1038).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=103908;
RN [1]_TaxID=103908;
RP SEQUENCE FROM N.A.
RA Bergelson J.M.;
RT "Receptor interactions, infectious cDNA, and nucleotide sequences of echovirus 1/8";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 69-330 FROM N.A.
RX MEDLINE=96203959; PubMed=8627260;
RT Huttunen P., Santti J., Pulli T., Hyypia T.;
RT "The major echovirus group is genetically coherent and related to coxsackie B viruses";
RL J. Gen. Virol. 77:715-725 (1996).
RN [3]
RP SEQUENCE OF 570-853 FROM N.A.
RX MEDLINE=99138973; PubMed=9971773;

RA Oberste M.S., Maher K., Kilpatrick D.R., Pallansch M.A.;
RT "Molecular evolution of the human enteroviruses: correlation of serotype with VP1 sequence and application to picornavirus classification.";
RL J. Virol. 73:1941-1948 (1999).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (3.55 ANGSTROMS) OF 2-850.
RX MEDLINE=99192672; PubMed=10089503;
RA Fillman D.J., Wien M.W., Cunningham J.A., Bergelson J.M., Hogle J.M.;
RT "Structure determination of echovirus 1";
RL Acta Crystallogr. D 54:1261-1272 (1998).
CC -!- FUNCTION: P2A AND THE P3C POLYPEPTIDES ARE PROTEASES THAT CLEAVE AT CERTAIN Q/G SITES IN THE POLYPEPTIDE. THEY ARE CYSTEINE PROTEASES.
CC -!- FUNCTION: It is thought that the P2C protein attaches to vesicular membranes and is associated with viral RNA synthesis.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the poliovirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Tyr-Gly bond in the picornavirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA) (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins. CLEAVAGE BETWEEN VP4 AND VP2 IS AUTOCATALYTIC; VP1/P2A IS CATALYZED BY P2A; ALL OTHER CLEAVAGES ARE CATALYZED BY P3C.
CC -!- SIMILARITY: P3C PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -!- SIMILARITY: P2A PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF029859; AAC63944.2; -.
CC EMBL; X89531; CAA61710.1; -.
CC EMBL; AF081314; AAD17718.1; -.
CC PDB; 1EV1; 27-JAN-99.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR000199; Pept_3C_picorn.
CC InterPro; IPR000081; Peptidase_C3.
CC InterPro; IPR003138; Pico_P1A.
CC InterPro; IPR002527; Pico_P2B.
CC InterPro; IPR001676; Rnv.
CC InterPro; IPR006005; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_PSVir.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00548; Cys-protease_3C; 1.
CC Pfam; PF02226; Pico_P1A; 1.
CC Pfam; PF00947; Pico_P2A; 1.
CC Pfam; PF01552; Pico_P2B; 1.
CC Pfam; PF00073; rhv_3.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICVIRUSNS.
CC ProDom; PD001125; Cys_protease_3C; 1.
CC ProDom; PD001306; Pico_P2A; 1.
CC ProDom; PD001274; Pico_P2B; 1.
CC SMART; SM00382; AAA; 1.
KW Polyprotein; Coat protein; Core protein; Transferase; Myristate;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure;
KW Lipoprotein.

FT CHAIN	2	69	COAT PROTEIN VP4.	FT TURN	338	341
FT CHAIN	70	330	COAT PROTEIN VP2.	FT TURN	345	346
FT CHAIN	331	569	COAT PROTEIN VP3.	FT STRAND	353	353
FT CHAIN	570	850	COAT PROTEIN VP1.	FT TURN	356	357
FT CHAIN	851	1000	PICORNAIN 2A.	FT STRAND	369	370
FT CHAIN	1001	1099	CORE PROTEIN P2B.	FT STRAND	372	372
FT CHAIN	1100	1428	CORE PROTEIN P2C.	FT TURN	373	373
FT CHAIN	1429	1517	CORE PROTEIN P3A.	FT HELIX	374	377
FT CHAIN	1518	1539	GENOME-LINKED PROTEIN VP6.	FT TURN	378	378
FT CHAIN	1540	1722	PICORNAIN 3C.	FT STRAND	381	382
FT CHAIN	1723	2184	RNA-DIRECTED RNA POLYMERASE.	FT HELIX	390	393
FT ACT SITE	1686	1686	N-myristoyl glycine (in host).	FT TURN	394	394
FT ACT SITE	1700	1700	PROTEASE 3C (POTENTIAL).	FT HELIX	397	399
FT CONFLICT	238	238	PROTEASE 3C (POTENTIAL).	FT STRAND	400	405
FT CONFLICT	611	611	N -> M (IN REF. 2).	FT TURN	406	406
FT CONFLICT	645	645	A -> R (IN REF. 2).	FT TURN	409	410
FT CONFLICT	724	724	A -> V (IN REF. 3).	FT STRAND	412	417
FT CONFLICT	774	774	Y -> F (IN REF. 3).	FT TURN	420	422
FT STRAND	4	7	V -> I (IN REF. 3).	FT HELIX	424	427
FT STRAND	26	29	A -> T (IN REF. 3).	FT TURN	428	428
FT HELIX	36	38		FT HELIX	430	435
FT STRAND	45	45		FT TURN	436	437
FT TURN	50	50		FT STRAND	438	442
FT HELIX	51	54		FT STRAND	445	452
FT STRAND	57	57				
FT TURN	63	64				
FT TURN	80	81				
FT STRAND	83	86				
FT STRAND	91	94				
FT STRAND	101	102				
FT HELIX	103	105				
FT TURN	113	115				
FT STRAND	123	123				
FT HELIX	126	128				
FT TURN	129	129				
FT STRAND	133	134				
FT STRAND	138	140				
FT TURN	142	143				
FT STRAND	147	151				
FT TURN	152	152				
FT HELIX	153	157				
FT HELIX	159	167				
FT STRAND	168	180				
FT TURN	185	186				
FT STRAND	188	197				
FT TURN	198	198				
FT STRAND	203	203				
FT TURN	206	207				
FT HELIX	212	214				
FT TURN	215	215				
FT STRAND	217	217				
FT TURN	218	219				
FT STRAND	220	220				
FT STRAND	222	223				
FT STRAND	225	225				
FT TURN	232	233				
FT STRAND	235	235				
FT STRAND	238	238				
FT TURN	239	241				
FT TURN	242	245				
FT STRAND	246	246				
FT HELIX	248	253				
FT STRAND	256	260				
FT TURN	261	263				
FT STRAND	266	271				
FT STRAND	280	280				
FT TURN	282	284				
FT STRAND	285	285				
FT STRAND	288	299				
FT TURN	302	303				
FT STRAND	308	324				

Query Match

Best Local Similarity 4.2%; Score 121.5; DB 1; Length 2184;

Matches 126; Conservative 73; Mismatches 178; Indels 339; Gaps 29;

QY	11	SSDGAAGLVPEINNEVM-----PLEPVAGASLATPV-----	41
Db	263	TNNSATIWPYVNSVPMNDNMYRHHNFTLMIIPFVFLDFSAGASTYVPTVTVAPMAEYN	322
QY	42	---VGQNI-----IDPWIRNNFVOAPAG--EFTVSPR-NSPGEILLLEL-----	81
Db	323	GURLAGHOGLPMTNTPGNSQFLTSDDFSPSAMPQFDVTPEMHIFGEVRNLMIEIAEVS	382
QY	82	-----GPDLPN-----YLAHLARMYNGHA	100
Db	383	MPINNDAAKVSSMEAYRVELSTNINAGTQVGFQNLFGAESVMNRITLMGEILNYVAHWS	442
QY	101	GMGEVOIVLAGNAFTAGKIIFAAIPGPPYENLSPQITMCPHVIIDVROLEPFLPMPD	160
Db	443	GSIKITFVFCGSAMTTGKLLSYAPPGAGPK-TRKDALMLGTHVYVMDVGLQSSCVLCIP-	500
QY	161	IWNFFHY-----NOG-----NDPKRLVAMLYTPLRANNGDDVF	196
Db	501	-WISQTHRFVEKOPYTNAGFVTCWYQTSVSPASNQPKCYMCMV-----SACNDF	551
QY	197	TVSCRVLTKPSDDEF-----TFLVPPTVE-SKTQKQFALPILKISEMTNSRPP	243
Db	552	SVRLMRDITKFIQTSTFYQGDVQNAVEGAMVRVADTVQTSATNSERVNPLTAVETGHTSQA	611
QY	244	VPVDVWYARNENQVVPQNGRVTLDGELLGTTPLLANVICXPKGEVTAQNG---DVRSY	300
Db	612	VPQDMQTRHVINNHVRSES---TTIEN-----FLARSACVFYLEY--KTGTKEDSNF	659
QY	301	-----RMDMETNTDGTPIDPTEDT-----	320
Db	660	NNWVITTRRVAQLRRKLEMFYVLRFDMEIT-----VWITSSQDQSTSQONAPVLTHQIMY	715
QY	321	-----PGPIGSPDF-----	332
Db	716	VPPGGPIPVSDVDYSQWTSNPSIFWTEGNAPARMSIPFISIGNAYSNFYDQWSHFSQAG	775
QY	333	LFGVASQRNK-----NEQNPAATRAHEAINTGGDH-----LCPOISSSEIYL	374
Db	776	VYGFITLNNMGOLFRRHVNKNPAAITSVARIYFKPKHVRVWVPRPPRLCPYINSTNV--	833
QY	375	TSPNILRCTNPQPLPQSGLRGTILIRSDNGHCHDMVGTSPTTTPWQQ-----	422
Db	834	-----NFEFKPVTEVRTNII-----TTGAFQOQSGAVYVGNRVV	868

QY 423 -----WRR-----SRGSN-----CCSSGHYPV 441
DB 869 NRHLATHIDWQNCWEDYNDRLVSTTARGCDTIARCOCTGTVFCLSRNKHYPV 924

RESULT 14
POLG HE71M STANDARD; PRT: 2193 AA.
AC Q66479;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VP3 (P2B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE Human enterovirus 71 (strain 7423/MS/87) (Ev 71).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.
OC
OX NCBI_TaxID=103922;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96434998; PubMed=8837884;
RA Brown B.A., Pallansch M.A.;
RT "Complete nucleotide sequence of enterovirus 71 is distinct from poliovirus."
RL Virus Res. 39:195-206(1995).
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the poliovirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA} (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; U22522; AAB39969.1; -.
CC HSSP; P03300; 1POV.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR004004; Calci_pol_hel.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000199; Pept_3C_picorn.
DR InterPro; IPR000081; Peptidase_C3.
DR InterPro; IPR003138; Pico_P2A.
DR InterPro; IPR002527; Pico_P2B.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_P5vir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00548; Cys-protease_3C; 1.
DR Pfam; PF02226; Pico_P1A; 1.
DR Pfam; PF00947; Pico_P2A; 1.
DR Pfam; PF01552; Pico_P2B; 1.
DR Pfam; PF00073; rhv; 3.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICVIRUSNS.
DR ProDom; PD001125; Cys_protease_3C; 1.

DR ProDom; PD001306; Pico_P2A; 1.
DR ProDom; PD001274; Pico_P2B; 1.
DR SMART; SM00382; AAA; 1.
KW Polyprotein; Coat protein; Core protein; Core proteinase; Thiol protease; Myristate;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
KW Lipoprotein.
FT CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 323 COAT PROTEIN VP2.
FT CHAIN 324 565 COAT PROTEIN VP3.
FT CHAIN 566 862 COAT PROTEIN VP1.
FT CHAIN 863 1012 CORE PROTEIN P2A.
FT CHAIN 1013 1111 CORE PROTEIN P2B.
FT CHAIN 1112 1440 CORE PROTEIN P2C.
FT CHAIN 1441 1526 CORE PROTEIN P3A.
FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1549 1731 PICORNAIN 3C.
FT CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE P3D.
FT LIPID 2 2 N-myristoyl glycine (in host) (By similarity).
FT ACT_SITE 1695 1695 PROTEASE (POTENTIAL).
FT ACT_SITE 1709 1709 PROTEASE (POTENTIAL).
SQ SEQUENCE 2193 AA; 242656 MW; 35E1B3CF88A50EF CRC64;
Query Match 4.1%; Score 120; DB 1; Length 2193;
Best Local Similarity 20.5%; Pred. No. 1;
Matches 115; Conservative 62; Mismatches 199; Indels 184; Gaps 27;
QY 58 QAPAGEFTVSPRNSPGEILLDELGLPDLNPYLALHARMYNGHAGGMEVQIVLAGNAPTAG 117
DB 399 QAGKSELCAVFRADFGD-----GFWQSTMLGQLCGYQWWSGLSEVTFMTGSPMATG 452
QY 118 KIIFRAIPGPGFYENLSPSQIT--MCPHVIVDROLEPFLPMPDINNNFFHYNNQNDPK 175
DB 453 KMLIAYTPRGGP---LPKDRATAMLTGTHVIMDFGLQSSVTLVIP--WISNTHY----- 500
QY 176 LRLVAMLYTPLRANSGDDVFTVSCRVLTKPSPDFEFTLVPTVESKTKQFALPKLS 235
DB 501 -----RA-HARDGVF-----DYTTGLV-----SIW 520
QY 236 EMTNSRFPVPVDMVTARNENQVQPQNGRVTLDGELGTTPLLAVNICKFGEVIKNG 295
DB 521 YQTN--YVVPICAPNTAYILALAAQKN-----FTMKLCKDTHIL-QTA 562
QY 296 DVRSYRMDMEINTDGTIDP--TEDTPGIPSPDQGLFGVAGRNKNEQNPAFRAHE 353
DB 563 SIQGRVADVIESSIGDSVSRALTQALPAPTQNT-----QVSSHRLDTGEVPALQAAE 616
QY 354 --AINTGCDHLCPOQISSSEIYLTSPNILRCTNPQPLPOSGLRGITLI-----RSDNGHC 406
DB 617 IGASNTSDSMI-----ETRCVLNHSHTAETLTDSFFSRAGLVGEIDLPLEGTNPNGYA 672
QY 407 H-----DMVGTSPPTPTWPO--OWRRCSRSGNSCCSSGH 437
DB 673 NWDIDITGYAQRKRVKVELFTYMRFAEFTFVACTPTGEVVPQLQLQWFMVPPGA----- 725
QY 438 RYFVPVVMNRVTWIVLSHKSGFSTTRKLPQLNLRWPLIRFINPDTGRVLFEARL----- 492
DB 726 --PKESRESLAWQTATNPSPVKLTDPDAQVS-----VPFMSASAYQWFDGYDTFGS 778
QY 493 HKQ-----GFTVAHTGDN-----PIVMPNGYFRFE---NWV-----NQPY 526
DB 779 HKQEKDLEYGACPNMMGTFSVRTVGSKSKYPLVV--RIYMRMGHVRAMIPRPMENQNY 836
QY 527 -----SLAPVGTGK 535
DB 837 LFKANPNYAGNSIKPTGTSR 856

RESULT 15
POLG FMDV1
ID POLG FMDV1 STANDARD; PRT: 2333 AA.
AC P03306; Q64768; Q84751; Q84752; Q84753; Q84754; Q84760; Q84761; Q84762; Q84763; Q84764; Q84765; Q84766; Q84767; Q84768;
ProDom; PD001125; Cys_protease_3C; 1.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:35:27 ; Search time 31.4978 Seconds

(without alignments)
5419.275 Million cell updates/sec

Title: US-09-926-799-11

Perfect score: 2894

Sequence: 1 MKMASNDAPSDGAGLVP.....VNOFYSLAPVGTGKRRRVQ 541

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mnc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2193.5	75.8	540	12 Q918B0	Q918b0 norwalk-lik
2	2179.5	75.3	540	12 Q91V49	Q91v49 human calic
3	1837	63.5	537	12 Q913B7	Q913b7 human calic
4	1827	63.1	537	12 Q91H09	Q91h09 human calic
5	1816.5	62.8	536	12 Q9BAC6	Q9bac6 human calic
6	1815	62.7	537	12 Q9V765	Q9v765 norwalk vir
7	1799	62.2	537	12 Q9PY74	Q9py74 human calic
8	1743	60.2	522	12 Q9QAX2	Q9qax2 norwalk-lik
9	1688.5	58.3	548	12 Q918A1	Q918a1 norwalk-lik
10	1688.5	58.3	548	12 Q90FK5	Q90fk5 human calic
11	1688.5	58.3	548	12 Q90FK3	Q90fk3 human calic
12	1687.5	58.3	548	12 Q91722	Q91722 norwalk-lik
13	1684.5	58.2	548	12 Q9V768	Q9v768 norwalk vir
14	1679.5	58.0	548	12 Q90FK9	Q90fk9 human calic
15	1679.5	58.0	548	12 Q90FK8	Q90fk8 human calic
16	1678.5	58.0	548	12 Q9BC87	Q9bc87 human calic

17	1678.5	58.0	548	12 Q90FL1	Q90fl1 human calic
18	1677.5	58.0	548	12 Q916E5	Q916e5 human calic
19	1677.5	58.0	548	12 Q90FK6	Q90fk6 human calic
20	1676.5	57.9	548	12 Q91725	Q91725 norwalk-lik
21	1675.5	57.9	548	12 Q90FK4	Q90fk4 human calic
22	1674.5	57.9	548	12 Q90FK1	Q90fk1 human calic
23	1673.5	57.8	548	12 Q9V0P5	Q9v0p5 human calic
24	1673.5	57.8	548	12 Q9V0P4	Q9v0p4 human calic
25	1672.5	57.8	548	12 Q66296	Q66296 calicivirus
26	1671.5	57.8	548	12 Q88291	Q88291 small round
27	1671.5	57.8	548	12 Q8V771	Q8v771 norwalk vir
28	1671.5	57.8	548	12 Q8BC90	Q8bc90 human calic
29	1671.5	57.8	548	12 Q90FK2	Q90fk2 human calic
30	1668.5	57.7	548	12 Q91728	Q91728 norwalk-lik
31	1668.5	57.7	548	12 Q915D3	Q915d3 human calic
32	1668.5	57.7	548	12 Q915D2	Q915d2 human calic
33	1667.5	57.6	548	12 Q91V40	Q91v40 human calic
34	1665.5	57.6	548	12 Q9PYA7	Q9pya7 human calic
35	1662.5	57.4	548	12 Q90FK7	Q90fk7 human calic
36	1662	57.4	535	12 Q911I5	Q911i5 human calic
37	1661.5	57.4	548	12 Q918A4	Q918a4 norwalk-lik
38	1659	57.3	535	12 Q915C6	Q915c6 human calic
39	1658.5	57.3	548	12 Q90FL2	Q90fl2 human calic
40	1657.5	57.3	548	12 Q68291	Q68291 human calic
41	1657	57.3	535	12 Q917X7	Q917x7 norwalk-lik
42	1657	57.3	535	12 Q8V784	Q8v784 norwalk vir
43	1656	57.2	535	12 Q9QMK6	Q9qmk6 chitta viru
44	1654.5	57.2	548	12 Q90FLO	Q90flo human calic
45	1653	57.1	535	12 Q68104	Q68104 hawaii calic

ALIGNMENTS

RESULT 1

Q918B0 ID Q918B0 PRELIMINARY; PRT; 540 AA.
AC Q918B0;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Gwynedd/273/1994/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171841;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Gwynedd/273/1994/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify a 3-kilobase region from the RNA polymerase gene to the poly(A) tail of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Gwynedd/273/1994/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Gwynedd/273/1994/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RT Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414409; AAL12980.1;
DR InterPro; IPR004005; Calici.coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calicicoat; 1.
SQ SEQUENCE 540 AA; 59384 MW; FD10B8F00CEAFF79 CRC64;

Query Match 75.8%; Score 2193.5; DB 12; Length 540;
Best Local Similarity 78.5%; Pred. No. 9.5e-183; Mismatches 26; Indels 17; Gaps 4;
Matches 431; Conservative 26; Mismatches 75; Indels 17; Gaps 4;

QY 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFFVQAP 60
DB 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFFVQAP 60
QY 61 AGEFTVSPRNSPGEILLDELGPDLNPYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120
DB 61 AGEFTVSPRNSPGEILLDELGPDLNPYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120
QY 121 FFAIIPGFPYENLSPQITMCPHVIIDVQLEPFLPMPDIWNPFHYNQNDPKRLVA 180
DB 121 FFAIIPGFPYENLSPQITMCPHVIIDVQLEPFLPMPDIWNPFHYNQNDPKRLVA 180
QY 181 MLYTPLRANSGDDVFTVSCRVLTKPSDPFEFTFLVPPTVESKTKQFALPILKISEMTNS 240
DB 181 MLYTPLRANSGDDVFTVSCRVLTKPSDPFEFTFLVPPTVESKTKQFALPILKISEMTNS 240
QY 241 RFPVPVDMYTARNENOVQVQNGRVTLDGELLGTTPLLA VNI CKFKGEVIAKNGDVRSY 300
DB 241 RFPVPVDMYTARNENOVQVQNGRVTLDGELLGTTPLLA VNI CKFKGEVIAKNGDVRSY 300
QY 301 RMDMETNTDGTPIIDPTEDTPGIPGSPDFQGIILFGVASORNKNEONPATRAHEAIINTGG 360
DB 301 RMDMETNTDGTPIIDPTEDTPGIPGSPDFQGIILFGVASORNKNEONPATRAHEAIINTGG 360
QY 361 DHLCPQIS-----SSEIYLTSPNLRCTNPQPLPQSLGRGTI-----LIRSDNGHCHDMVGT 413
DB 361 DHLCPQIS-----SSEIYLTSPNLRCTNPQPLPQSLGRGTI-----LIRSDNGHCHDMVGT 413
QY 414 P-TTPTWQVQWRRRCRGSCNCCSGHRYPPVPMNRVTWIVLSHKSGFSTSTRKLPQNLRL 472
DB 414 P-TTPTWQVQWRRRCRGSCNCCSGHRYPPVPMNRVTWIVLSHKSGFSTSTRKLPQNLRL 472
QY 420 PAVAPLFPGEQILFFRSQIPSSGGHELGYMDCLVPOEWQHFYQEAATAQSEV----- 472
DB 420 PAVAPLFPGEQILFFRSQIPSSGGHELGYMDCLVPOEWQHFYQEAATAQSEV----- 472
QY 473 WPLIRFNPDTGRVLFPEARLHKQGITVAHTGDNPIVMPNGYFRFEAWNQFYSLAPVG 532
DB 473 WPLIRFNPDTGRVLFPEARLHKQGITVAHTGDNPIVMPNGYFRFEAWNQFYSLAPVG 532
QY 533 TGKGRRRVQ 541
DB 533 TGKGRRRVQ 540

RESULT 2
Q91V49 PRELIMINARY; PRT; 540 AA.
AC Q91V49; PRELIMINARY; PRT; 540 AA.
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/MLV/Leeds/90/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=122912;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/MLV/Leeds/90/UK;
RX MEDLINE=20404883; PubMed=10949950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL; AJ276608; CAB89089.1;
DR InterPro; IPR004005; Calici coat.
DR InterPro; IPR008975; Viral gap coat.
DR Pfam; PF00915; Calici coat; 1.
SQ SEQUENCE 540 AA; 59374 MW; 5FC8FBDCBDC8DAD3 CRC64;

Query Match 75.3%; Score 2179.5; DB 12; Length 540;

Best Local Similarity 78.5%; Pred. No. 1.6e-181;
Matches 431; Conservative 23; Mismatches 78; Indels 17; Gaps 4;

QY 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFFVQAP 60
DB 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFFVQAP 60
QY 61 AGEFTVSPRNSPGEILLDELGPDLNPYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120
DB 61 AGEFTVSPRNSPGEILLDELGPDLNPYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120
QY 121 FFAIIPGFPYENLSPQITMCPHVIIDVQLEPFLPMPDIWNPFHYNQNDPKRLVA 180
DB 121 FFAIIPGFPYENLSPQITMCPHVIIDVQLEPFLPMPDIWNPFHYNQNDPKRLVA 180
QY 181 MLYTPLRANSGDDVFTVSCRVLTKPSDPFEFTFLVPPTVESKTKQFALPILKISEMTNS 240
DB 181 MLYTPLRANSGDDVFTVSCRVLTKPSDPFEFTFLVPPTVESKTKQFALPILKISEMTNS 240
QY 241 RFPVPVDMYTARNENOVQVQNGRVTLDGELLGTTPLLA VNI CKFKGEVIAKNGDVRSY 300
DB 241 RFPVPVDMYTARNENOVQVQNGRVTLDGELLGTTPLLA VNI CKFKGEVIAKNGDVRSY 300
QY 301 RMDMETNTDGTPIIDPTEDTPGIPGSPDFQGIILFGVASORNKNEONPATRAHEAIINTGG 360
DB 301 RMDMETNTDGTPIIDPTEDTPGIPGSPDFQGIILFGVASORNKNEONPATRAHEAIINTGG 360
QY 361 DHLCPQIS-----SSEIYLTSPNLRCTNPQPLPQSLGRGTI-----LIRSDNGHCHDMVGT 413
DB 361 DHLCPQIS-----SSEIYLTSPNLRCTNPQPLPQSLGRGTI-----LIRSDNGHCHDMVGT 413
QY 414 P-TTPTWQVQWRRRCRGSCNCCSGHRYPPVPMNRVTWIVLSHKSGFSTSTRKLPQNLRL 472
DB 414 P-TTPTWQVQWRRRCRGSCNCCSGHRYPPVPMNRVTWIVLSHKSGFSTSTRKLPQNLRL 472
QY 420 PAVAPLFPGEQILFFRSQIPSSGGHELGYMDCLVPOEWQHFYQEAATAQSEV----- 472
DB 420 PAVAPLFPGEQILFFRSQIPSSGGHELGYMDCLVPOEWQHFYQEAATAQSEV----- 472
QY 473 WPLIRFNPDTGRVLFPEARLHKQGITVAHTGDNPIVMPNGYFRFEAWNQFYSLAPVG 532
DB 473 WPLIRFNPDTGRVLFPEARLHKQGITVAHTGDNPIVMPNGYFRFEAWNQFYSLAPVG 532
QY 533 TGKGRRRVQ 541
DB 533 TGKGRRRVQ 540

RESULT 3
Q913B7 PRELIMINARY; PRT; 537 AA.
AC Q913B7; PRELIMINARY; PRT; 537 AA.
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/Idaho Falls/378/1996/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=173723;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/Idaho Falls/378/1996/US;
RX MEDLINE=99034638; PubMed=9815206;
RA Fankhauser R.L., Noel J.S., Monroe S.S., Ando T., Glass R.I.;
RT "Molecular epidemiology of 'Norwalk-like' viruses' in outbreaks of
RT gastroenteritis in the United States.";
RL J. Infect. Dis. 178:1571-1578(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/Idaho Falls/378/1996/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like' viruses'.";
RL J. Infect. Dis. 181:8336-8348(2000).
RN [3]
RP SEQUENCE FROM N.A.

RC STRAIN=NLV/Idaho Falls/378/1996/US;
 RA Noel J.S., Fankhauser R.L., Monroe S.S., Glass R.I., Ando T.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY054299; AAL15547.1; --
 DR InterPro; IPR004005; Calici_coat.
 DR Pfam; PF00915; Calici_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 SQ SEQUENCE 537 AA; 58600 MW; 70ABEACFC1F5E05D CRC64;

Query Match 63.5%; Score 1837; DB 12; Length 537;
 Best Local Similarity 64.5%; Pred. No. 1.3e-151;
 Matches 362; Conservative 57; Mismatches 98; Indels 44; Gaps 9;

QY 1 MKVASNDAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVQQNIIDPWIRNNFVQAP 60
 DB 1 MKVASNDAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVQQNIIDPWIRNNFVQAP 60
 QY 61 AGEFTVSPRNSPGEILLDELGLDLPNLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120
 DB 61 AGEFTVSPRNSPGEILLDELGLDLPNLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120
 QY 121 FAAIPGPPFYENLSQITMCPHVIVDVRLQLEPFLPMPDIWNNFPHYNGNDPKRLIVA 180
 DB 121 FAAIPGPPFYENLSQITMCPHVIVDVRLQLEPFLPMPDIWNNFPHYNGNDPKRLIVA 180
 QY 181 MLYTPLRANNSGDDVFTVSCRVLTKPSPDFEFTFLVPPVSKTKQFALPILKISEMTNS 240
 DB 181 MLYTPLRANNSGDDVFTVSCRVLTKPSPDFEFTFLVPPVSKTKQFALPILKISEMTNS 240
 QY 241 REPVPVDMYTARNENQVVPQNGRVTLDGELLGTTPLLVANNICKPKGEVIAK-NGDVRS 299
 DB 241 REPVPVDMYTARNENQVVPQNGRVTLDGELLGTTPLLVANNICKPKGEVIAK-NGDVRS 299
 QY 300 YRMDMEITNDGTPTDPTEDTGPISGDPFGILFGVASQRNKNQONPATRAHEAIIING 359
 DB 300 YRMDMEITNDGTPTDPTEDTGPISGDPFGILFGVASQRNKNQONPATRAHEAIIING 359
 QY 301 YQ--LQTLNLDGSPIDTDDTGPGLCPDFTGLLYGVASQRPD---ATRAHEAIIING 355
 DB 301 YQ--LQTLNLDGSPIDTDDTGPGLCPDFTGLLYGVASQRPD---ATRAHEAIIING 355
 QY 360 GDHLCPQISSSIYLTSPNLRCTNPQPLPQSLRGTLIRSDNGCHDMVGTSPPTPTW 419
 DB 360 GDHLCPQISSSIYLTSPNLRCTNPQPLPQSLRGTLIRSDNGCHDMVGTSPPTPTW 419
 QY 356 SDTAPKIGQVRFYSTSSNFER-NQTHFTPIGI-----YIEGNSDF----- 397
 DB 356 SDTAPKIGQVRFYSTSSNFER-NQTHFTPIGI-----YIEGNSDF----- 397
 QY 420 PQWRCRSGSCCCSGHRYPV--PVMNRVTWIVLSHKSGFSTSTR-----KLPO---- 468
 DB 398 -NQWQLPRYGGHLANNHLLAPAVSLPFGCEQLFRFSFIPGASGHTNGEMDCLLPQEFVQ 456
 QY 469 -----LNLRWPLIRFNPTDGRVLPFARLHKQGFITVAHTGDNPIVMPNGYFRPEA 520
 DB 457 HFYQEAATAQSEVALLRFVNPDTGRVLPFESKLHKQGFMTIASSGDHPIIMPTNGYFRPEA 516
 QY 521 WYNQFYSAPVGTGKRRRVQ 541
 DB 517 WYNQFYSAPVGTGSGRRRIQ 537

RESULT 4
 Q91H09 ID Q91H09 PRELIMINARY; PRT; 537 AA.
 AC Q91H09
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DE Major capsid protein VPI.
 OS Human calicivirus NLV/VA97207/1997.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 OC NCBI_TaxID=165510;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NLV/VA97207/1997;
 RA Jiang X., Zhong W.M., Farkas T., Huang P.W., Wilton N., Barrett E.,
 RA Fulcon D., Morrow R., Matson D.O.;
 RT "Baculovirus expression and antigenic characterization of the capsid
 RT proteins of three Norwalk-like viruses.";

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY038599; AAK84676.1; --
 DR InterPro; IPR004005; Calici_coat.
 DR Pfam; PF00915; Calici_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 SQ SEQUENCE 537 AA; 58588 MW; E6D1B1F858CA0EA7 CRC64;

Query Match 63.1%; Score 1827; DB 12; Length 537;
 Best Local Similarity 64.2%; Pred. No. 9.8e-151;
 Matches 360; Conservative 56; Mismatches 101; Indels 44; Gaps 9;

QY 1 MKVASNDAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVQQNIIDPWIRNNFVQAP 60
 DB 1 MKVASNDAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVQQNIIDPWIRNNFVQAP 60
 QY 61 AGEFTVSPRNSPGEILLDELGLDLPNLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120
 DB 61 AGEFTVSPRNSPGEILLDELGLDLPNLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120
 QY 121 FAAIPGPPFYENLSQITMCPHVIVDVRLQLEPFLPMPDIWNNFPHYNGNDPKRLIVA 180
 DB 121 FAAIPGPPFYENLSQITMCPHVIVDVRLQLEPFLPMPDIWNNFPHYNGNDPKRLIVA 180
 QY 181 MLYTPLRANNSGDDVFTVSCRVLTKPSPDFEFTFLVPPVSKTKQFALPILKISEMTNS 240
 DB 181 MLYTPLRANNSGDDVFTVSCRVLTKPSPDFEFTFLVPPVSKTKQFALPILKISEMTNS 240
 QY 241 REPVPVDMYTARNENQVVPQNGRVTLDGELLGTTPLLVANNICKPKGEVIAK-NGDVRS 299
 DB 241 REPVPVDMYTARNENQVVPQNGRVTLDGELLGTTPLLVANNICKPKGEVIAK-NGDVRS 299
 QY 300 YRMDMEITNDGTPTDPTEDTGPISGDPFGILFGVASQRNKNQONPATRAHEAIIING 359
 DB 300 YRMDMEITNDGTPTDPTEDTGPISGDPFGILFGVASQRNKNQONPATRAHEAIIING 359
 QY 301 YQ--LQTLNLDGSPIDTDDTGPGLCPDFTGLLYGVASQRPD---ATRAHEAIIING 355
 DB 301 YQ--LQTLNLDGSPIDTDDTGPGLCPDFTGLLYGVASQRPD---ATRAHEAIIING 355
 QY 360 GDHLCPQISSSIYLTSPNLRCTNPQPLPQSLRGTLIRSDNGCHDMVGTSPPTPTW 419
 DB 360 GDHLCPQISSSIYLTSPNLRCTNPQPLPQSLRGTLIRSDNGCHDMVGTSPPTPTW 419
 QY 356 SDTAPKIGQVRFYSTSSN-FETNQTHFTPIGI-----YIEGNSDF----- 397
 DB 356 SDTAPKIGQVRFYSTSSN-FETNQTHFTPIGI-----YIEGNSDF----- 397
 QY 420 PQWRCRSGSCCCSGHRYPV--PVMNRVTWIVLSHKSGFSTSTR-----KLPO---- 468
 DB 398 -NQWQLPRYGGHLANNHLLAPAVSLPFRQQLFRFSFIPGASGHTNGEMDCLLPQEFVQ 456
 QY 469 -----LNLRWPLIRFNPTDGRVLPFARLHKQGFITVAHTGDNPIVMPNGYFRPEA 520
 DB 457 HFYQEAATAQSEVALLRFVNPDTGRVLPFESKLHKQGFMTIASSGDHPIIMPTNGYFRPEA 516
 QY 521 WYNQFYSAPVGTGKRRRVQ 541
 DB 517 WYNQFYSAPVGTGSGRRRIQ 537

RESULT 5
 Q9BAC6 ID Q9BAC6 PRELIMINARY; PRT; 536 AA.
 AC Q9BAC6
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Major capsid protein VPI.
 OS Human calicivirus NLV/M7/1999/US.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Norovirus.
 OC NCBI_TaxID=207241;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Hu/NLV/M7/1999/US;
 RA Vinje J.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY130761; AAN05735.1; --
 DR InterPro; IPR004005; Calici_coat.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00915; Calici_coat; 1.

```
SQ SEQUENCE 536 AA; 58883 MW; C3BDC61967FC8765 CRC64;
Query Match 62.8%; Score 1816.5; DB 12; Length 536;
Best Local Similarity 64.3%; Pred. No. 8.1e-150;
Matches 367; Conservative 52; Mismatches 87; Indels 65; Gaps 12;

QY 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGOQNIIDPWIRNNFVQAP 60
DB 1 MKMASNDATPSDGAAGLVPEINNEVMALEPVAGASIAAPVVGQNIIDPWIRNNFVQAP 60

QY 61 AGEFTVSPRNSPGEILLDLELGPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120
DB 61 AGEFTVSPRNSPGEILLDLELGPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKIL 120

QY 121 FFAIPGPFYENLSQITMCPHVIIIDVQLPEFLLPMPDIWNNFFHYNQNDPKRLVA 180
DB 121 FFAIPGPFYENLSQAQITMCPHVIIIDVQLPEFLLPMPDIRNVFYHYNQNSPKRLVA 180

QY 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDPFETFLVPPTVESKTKQFALPILKISEMTNS 240
DB 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDPFETFLVPPTVESKTKNFTLPVLURVSEMTNS 240

QY 241 RFPFVDMYTARNENQVQPQNGRVTLTGELGTTPLLAVNICKFKGEVIARNGDVRYSY 300
DB 241 RFPFVLDQMYTSRNNENIIVQPQNGRCTTDGELLGTTTLOSVCINFRGTMQAKLNEQPRY 300

QY 301 RMDMEITNDGTPIDTPTDTPGIGSPDQFQILFGVASORNKNEQNPAATRAHEAINTGG 360
DB 301 Q--LQLTNLDGSPIDTDDMPAPLGPDPQAMLYGVASQSRSDN--ATRAHDAQIDTAG 356

QY 361 DHLCPQI-----SSSEIYLTSPNLRCTNPOPLQPSGLRGTLIRNSNGHCHDMVGTS 413
DB 357 DTFAPKIGQVRKSSDDDFDLHDP-----TKFTP-----IGNVDDQH----- 394

QY 414 PTTPTWP--QOWRRCRSGSCSSGHRYP--VPVVMNRVTWIVLSH--KSGFSTSTR---K 465
DB 395 -----PFRQWSLPNYGGHLALNHLAPAVTLPFGCEQLFFRSHPISAGGHTDGAIDCL 448

QY 466 LPQMLRW-----PLIRFINDTGRVLFEARLHKQGFIVAHGTGNPIVM 510
DB 449 LPQ---EWIEHFVQEAPSQSDIALVRFINPDTGRVLFEARLHKQGFIVAHGTGNPIVM 505

QY 511 PNGVYFRFAWVNOFYSLAPVGTGKRRRVQ 541
DB 506 PINGVFRFAWVNPFTYLPVGTGSRRIQ 536

RESULT 6
Q8V765
ID Q8V765 PRELIMINARY; PRT; 537 AA.
AC Q8V765;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF2 protein (Capsid).
OS Norwalk virus, and
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=11983, 95340;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Norwalk virus; STRAIN=U25GII;
RA Kojima S., Kageyama T., Fukushi S., Hoshino F.B., Shinohara M.,
RA Uchida K., Natori K., Takeda N., Katayama K.;
RT "Genogroup-specific primers for detect Norwalk like virus.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Norwalk-like virus; STRAIN=Saitama U25;
RA Katayama K.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
```

```
RP SEQUENCE FROM N.A.
RC SPECIES=Norwalk-like virus; STRAIN=Saitama U25;
RX MEDLINE=22192455; PubMed=1202225;
RA Katayama K., Shirato-Horikoshi H., Kojima S., Kageyama T., Oka T.,
RA Hoshino F.B., Fukushi S., Shinohara M., Uchida K., Suzuki Y.,
RA Gojobori T., Takeda N.;
RT "Phylogenetic Analysis of the Complete Genome of 18 Norwalk-like
RT Viruses.";
RL Virology 299:225-239(2002).
DB EMBL; AB067543; BAB84158.1; --
DB EMBL; AB067543; BAC11831.1; --
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 537 AA; 58969 MW; 605A68451D0308B7 CRC64;

Query Match 62.7%; Score 1815; DB 12; Length 537;
Best Local Similarity 64.0%; Pred. No. 1.1e-149;
Matches 359; Conservative 59; Mismatches 99; Indels 44; Gaps 7;

QY 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGOQNIIDPWIRNNFVQAP 60
DB 1 MKMASNDAAAPSSDGAAGLVPEINNEVMAIEPVAGASLAAPVVGQNIIDPWIRNNFVQAP 60

QY 61 AGEFTVSPRNSPGEILLDLELGPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120
DB 61 AGEFTVSPRNSPGEILLDLELGPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKIL 120

QY 121 FFAIPGPFYENLSQITMCPHVIIIDVQLPEFLLPMPDIWNNFFHYNQNDPKRLVA 180
DB 121 FFAIPGPFYENLSQAQITMCPHVIIIDVQLPEFLLPMPDIRNTFFHYNQNSGPKRLVA 180

QY 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDPFETFLVPPTVESKTKQFALPILKISEMTNS 240
DB 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDPFETFLVPPTVESKTKAFTLPILKISEMTNS 240

QY 241 RFPFVDMYTARNENQVQPQNGRVTLTGELGTTPLLAVNICKFKGEVIARNGDVRYSY 300
DB 241 RFPFVDMYTARNENIIVQPQNGRVTLTGELGTTTLPVSICGFRGTLQTLRADQPNY 300

QY 301 RMDMEITNDGTPIDTPTDTPGIGSPDQFQILFGVASORNKNEQNPAATRAHEAINTGG 360
DB 301 TYQVHLENLDGSPVDTPDEVPAPLGPDPQALFGVISOQRSDN---ATRAHEARVNTND 357

QY 361 DHLCPQIS-----SSSEIYLTSPNLRCT-----NPQPLQPSGLRGTLIRNS 401
DB 358 PTFAPQIAQVRKPSNSDFFDNEP--IKFTPVGISVDSONSYNQWLLPRYG----- 406

QY 402 DNGHCHDMVGTSPT--TPTWQQWRRCRSGSCSSGHRYPVPVVMNRVTWIVLSHSGFS 460
DB 407 --GHLTNTHLAPSVSPWPGCEQLFFRSFWPGASGHTDGAIDCLLPQBWVAHFYQEAT 464

QY 461 TSTRKLPQLNLRWPLIRFINDTGRVLFEARLHKQGFIVAHGTGNPIVMPNGYFRFEA 520
DB 465 AQTDV-----ALIRFVNPDTRVLFEGLHKGQGITISNGDHPIMVMPANGYFRFEA 516

QY 521 WVNQFYSLAPVGTGKRRRVQ 541
DB 517 WVNQFYSLAPVGTGSRRIQ 537

RESULT 7
Q9PY74
ID Q9PY74 PRELIMINARY; PRT; 537 AA.
AC Q9PY74;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein (Fragment).
OS Human calicivirus strain Hu/NLW/Amsterdam/98-18/1998/NET.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=106515;
```



```

RN  SEQUENCE FROM N.A.
RP  STRAIN=Hu/NLIV/Amsterdam/98-18/1998/NET;
RX  MEDLINE=203411555; PubMed=10878050;
RA  Vinje J., Koopmans M.P.G.;
RT  "Simultaneous detection and genotyping of 'Norwalk-like viruses' by
RT  oligonucleotide array in a reverse line blot hybridization format.";
RL  J. Clin. Microbiol. 38:2595-2601(2000).
DR  EMBL; AF195848; AAF05820.1; -.
DR  InterPro; IPR004005; Calici coat.
DR  InterPro; IPR008975; Viral Cap_coat.
DR  Pfam; PF00915; Calici coat; 1.
FT  NON_TER 537
SQ  SEQUENCE 537 AA; 5891 MW; D8B6DD65CCA07C06 CRC64;

Query Match      62.2%; Score 1799; DB 12; Length 537;
Best Local Similarity 64.2%; Pred. No. 2.7e-148;
Matches 357; Conservative 64; Mismatches 101; Indels 34; Gaps 10;

QY  1 MKQASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNNFVQAP 60
DB  1 MKQASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNNFVQAP 60
QY  61 AGEFTVSPNSPGEILLDLELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120
DB  61 AGEFTVSPNSPGEILLDLELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120
QY  121 FFAIIPGFFPYENLSPQITMCPHVIIDVRLQEPFLPMPDIWNNEFHYNGNDPKRLVA 180
DB  121 FFAIIPGFFPYENLSPQITMCPHVIIDVRLQEPFLPMPDIWNNEFHYNGNDPKRLVA 180
QY  181 MLYTPLRANNSGDVFTVSCRVLTKPSDFEFTFLVPPTVESKTKQFALPILKISEMTNS 240
DB  181 MLYTPLRANNSGDVFTVSCRVLTKPSDFEFTFLVPPTVESKTKQFALPILKISEMTNS 240
QY  241 RPPVVDVMTARNENOVQPNQGRVTLDELGLGTTPLLAQVLSQSSDN---ATRAHEARVNTND 300
DB  241 RPPVVDVMTARNENOVQPNQGRVTLDELGLGTTPLLAQVLSQSSDN---ATRAHEARVNTND 300
QY  301 RMDMEITNDTPIDPTDTPGIPGSPDFQGLIFGVASQNKNEQNPAHAEAIINTGG 360
DB  301 TQVHLENLDGSPVDTDEVPAPLGTDPFQAQLFGVISQSSDN---ATRAHEARVNTND 357
QY  361 DHLCPQISSEIYLTSPNIRCTNPQL-----PQSLRGITLIRSDNGHCHDMV 410
DB  358 PTFAPQI--AQVREKSPSTDFSDN-EPIKFTPVGISVDSQNSYNQWLLPRY--GGHLTNT 413
QY  411 GTSPT-TPTWPOQ----WRCRSGNSCCSGHRYFPVPMNRVTWIVLSHKGSGFSTSTRK 465
DB  414 HLAPSVSPMFPGEQILFFRSFMPGAGGFTDG---AIDCLLPQ-EWVAHFYQEAATAQTDV 469
QY  466 LPQLNLRLPFRINPDTGRVLFEARLHKQGITVAHTGDNPIVMPNPGYFRFEAWVNOF 525
DB  470 -----ALIRFVNPDTGRVLFEGLKHKQGITISNSGDHPVIMPANGYFRFEAWVNOF 521
QY  526 YSLAPVGTGKRRRVQ 541
DB  522 YSLAPVGTGKRRRVQ 537

RESULT 8
Q9QAX2 PRELIMINARY; PRT; 522 AA.
AC Q9QAX2;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Capsid protein (Fragment).
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
RN [1]

```

```

RP  SEQUENCE FROM N.A.
RC  STRAIN=NLV114;
RX  MEDLINE=20150145; PubMed=10686033;
RA  Yoda T., Terano Y., Shimada A., Suzuki Y., Yamazaki K., Sakon N.,
RA  Oishi I., Utagawa E.T., Okuno Y., Shibata T.;
RT  "Expression of recombinant Norwalk-like virus capsid proteins using a
RT  bacterial system and the development of its immunologic detection.";
RL  J. Med. Virol. 60:475-481(2000).
DR  EMBL; AB028246; BAA89030.1; -.
DR  InterPro; IPR004005; Calici coat.
DR  InterPro; IPR008975; Viral Cap_coat.
DR  Pfam; PF00915; Calici coat; 1.
FT  NON_TER 522
SQ  SEQUENCE 522 AA; 57276 MW; 814CFD6CC4DF14A8 CRC64;

Query Match      60.2%; Score 1743; DB 12; Length 522;
Best Local Similarity 63.2%; Pred. No. 2.1e-143;
Matches 345; Conservative 59; Mismatches 98; Indels 44; Gaps 7;

QY  9 AFPSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNNFVQAPAGFTVSP 68
DB  1 AFPSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNNFVQAPAGFTVSP 60
QY  69 RNSPGEILLDLELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNAFTAGKIIIPAAIPPGF 128
DB  61 RNAPGEILLDLELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNAFTAGKIIIPAAIPPGF 120
QY  129 PYENLSPQITMCPHVIIDVRLQEPFLPMPDIWNNEFHYNGNDPKRLVAMLYTPLRA 188
DB  121 PYENLSPQITMCPHVIIDVRLQEPFLPMPDIWNNEFHYNGNDPKRLVAMLYTPLRA 180
QY  189 NNSGDVFTVSCRVLTKPSDFEFTFLVPPTVESKTKQFALPILKISEMTNSRFPVVDV 248
DB  181 NNAGEDVFTVSCRVLTKPSDFEFTFLVPPTVESKTKQFALPILKISEMTNSRFPVVDV 240
QY  249 MYTARNENOVQPNQGRVTLDELGLGTTPLLAQVLSQSSDN---ATRAHEARVNTND 308
DB  241 LYTSRNIIVQPNQGRVTLDELGLGTTPLLAQVLSQSSDN---ATRAHEARVNTND 300
QY  309 TGCTPIDTDETPGIPGSPDFQGLIFGVASQNKNEQNPAHAEAIINTGGHLCPOIS 368
DB  301 LDGSPVDTDEVPAPLGTDPFQAQLFGVISQSSDN---ATRAHEARVNTNDPTFAPQIA 357
QY  369 -----SSEIYLTSPNIRCTNPQL-----NPQPLPQSLRGITLIRSDNGHCHDM 409
DB  358 QVRFKSPSTDFSDN-EPIKFTPVGISVDSQNSYNQWLLPRY-----GHLTNN 404
QY  410 VGTSP-TPTWPOQWRCRSGNSCCSGHRYFPVPMNRVTWIVLSHKGSGFSTSTRKLPQ 468
DB  405 THLAPSVSPMFPGEQILFFRSFMPGAGGFTDG---AIDCLLPQ-EWVAHFYQEAATAQTDV 461
QY  469 LNLRLPFRINPDTGRVLFEARLHKQGITVAHTGDNPIVMPNPGYFRFEAWVNOF 528
DB  462 -----ALIRFVNPDTGRVLFEGLKHKQGITISNSGDHPVIMPANGYFRFEAWVNOF 516
QY  529 APVGTG 534
DB  517 APVGTG 522

RESULT 9
Q918A1 PRELIMINARY; PRT; 548 AA.
AC Q918A1;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Capsid protein.
OS ORF2.
OS Norwalk-like virus NLV/New Orleans/279/1994/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.

```

```

OX NCBI_TaxID=171844;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/279/1994/US;
RA MEDLINE=97193806; PubMed=9041391;
RX Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RA "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577 (1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/279/1994/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
RA Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995.";
RL J. Med. Virol. 53:372-383 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/279/1994/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.';
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/279/1994/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF14412; AAL12989.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 59990 MW; F998C603777F8595 CRC64;

Query Match 58.3%; Score 1688.5; DB 12; Length 548;
Best Local Similarity 58.4%; Pred. No. 1.3e-138;
Matches 333; Conservative 64; Mismatches 126; Indels 43; Gaps 6;

QY 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVQQNIIDPWIRNPFVQAP 60
DB 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVQQNIIDPWIRNPFVQAP 60
QY 61 AGEFTVSPRNSPGEIILLDELGPDLNPYLALARMYNGHAGMEVQIVLAGNAFTAGKII 120
DB 61 GGEFTVSPRNSPGEVLLNLELGPDLNPYLALARMYNGHAGMEVQIVLAGNAFTAGKII 120
QY 121 FFAIIPGPPYENLSPQITMCPHVIIDVRQLEPFLPMPDIWNNFFHYNQNDPKRLVA 180
DB 121 FFAIIPGPPYENLSPQITMCPHVIIDVRQLEPFLPMPDIWNNFFHYNQNDPKRLVA 180
QY 181 MLYTPLRANSGDDVFTVSCRVLTKPSDPFEFTFLVPPTVESKTKQFALPKIKISMTNS 240
DB 181 MLYTPLRANSGDDVFTVSCRVLTKPSDPFEFTFLVPPTVESKTKQFALPKIKISMTNS 240
QY 241 RFPVPVDMVYARNNOVQVQNGRVTLDGELGTTPLLAVNICKFGEVIAKNGDVR-- 298
DB 241 RFPVPVDMVYARNNOVQVQNGRVTLDGELGTTPLLAVNICKFGEVIAKNGDVR-- 298
QY 299 -----SYRMDMEITNTDGTPIIDPTEDTPGICSPDFOGILFGVASORNKNEQNAT 349
DB 299 -----SYRMDMEITNTDGTPIIDPTEDTPGICSPDFOGILFGVASORNKNEQNAT 349
QY 301 ADTPTPLRNFYWHIQLDNLNGTTPYDPAEDIPAGLTPDFRGKIVFGVASQRNPD--TT 357
DB 301 ADTPTPLRNFYWHIQLDNLNGTTPYDPAEDIPAGLTPDFRGKIVFGVASQRNPD--TT 357
QY 350 RAHEAIINTGGDHLCPQISSEIYLT-----PNILRCTNPQPLP-----QSLRG 395
DB 350 RAHEAIINTGGDHLCPQISSEIYLT-----PNILRCTNPQPLP-----QSLRG 395
QY 358 RAHEAKVDITSGRTFKLGSLEISTESDDFDN-----OPTKFTPVGVDNEADFOQ 410
DB 358 RAHEAKVDITSGRTFKLGSLEISTESDDFDN-----OPTKFTPVGVDNEADFOQ 410
QY 396 TILIRSDNGCHDMVGTSPPTPTWPOQWRRRCSCGNSCCSGHRYPPVPMNVRTWVLSH 455
DB 396 TILIRSDNGCHDMVGTSPPTPTWPOQWRRRCSCGNSCCSGHRYPPVPMNVRTWVLSH 455

Db 411 WSLPNYSQGFTHNNVLAPAVAPNPPGBQLLFFRSQRPSSGGRSNGILDCLVPQEWVQHIFY 470
QY 456 KSGFSTSTRKLPQLNLRWPLIRFNPDTGRVLFPEARLHKQGFITVAHTGDNPIVMPNGY 515
DB 471 QESAPAQTV-----ALVRYNPDTGRVLFPEARLHKLGFMTIAKNGDSPIVPPNGY 522
QY 516 FRFEAWNQFYSLAPVGTGKGRRRVQ 541
DB 523 FRFESWNVNPFYTLAPMGTCNGRRRIQ 548

RESULT 10
Q80FKS
ID Q80FKS PRELIMINARY; PRT; 548 AA.
AC Q80FKS;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/500-01/SWE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus
OX NCBI_TaxID=227521;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/500-01;
RA Nilsson M., Hedlund K.-O., Thorhagen M., Johansen K., Ekspong A.,
RA Svensson L.;
RT "Evolution of human calicivirus RNA in vivo: accumulation of mutations
RT in the protruding P2 domain of the capsid leads to structural changes
RT and possibly a new phenotype.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV247438; AAP03036.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 59856 MW; 34BC961A841B6EB8 CRC64;

Query Match 58.3%; Score 1688.5; DB 12; Length 548;
Best Local Similarity 59.4%; Pred. No. 1.3e-138;
Matches 337; Conservative 67; Mismatches 118; Indels 45; Gaps 8;

QY 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVQQNIIDPWIRNPFVQAP 60
DB 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVQQNIIDPWIRNPFVQAP 60
QY 61 AGEFTVSPRNSPGEIILLDELGPDLNPYLALARMYNGHAGMEVQIVLAGNAFTAGKII 120
DB 61 GGEFTVSPRNSPGEVLLNLELGPDLNPYLALARMYNGHAGMEVQIVLAGNAFTAGKII 120
QY 121 FFAIIPGPPYENLSPQITMCPHVIIDVRQLEPFLPMPDIWNNFFHYNQNDPKRLVA 180
DB 121 FFAIIPGPPYENLSPQITMCPHVIIDVRQLEPFLPMPDIWNNFFHYNQNDPKRLVA 180
QY 181 MLYTPLRANSGDDVFTVSCRVLTKPSDPFEFTFLVPPTVESKTKQFALPKIKISMTNS 240
DB 181 MLYTPLRANSGDDVFTVSCRVLTKPSDPFEFTFLVPPTVESKTKQFALPKIKISMTNS 240
QY 241 RFPVPVDMVYARNNOVQVQNGRVTLDGELGTTPLLAVNICKFGEVIAKNGDVR-- 296
DB 241 RFPVPVDMVYARNNOVQVQNGRVTLDGELGTTPLLAVNICKFGEVIAKNGDVR-- 296
QY 297 -----VRS--YMDMEITNTDGTPIIDPTEDTPGICSPDFOGILFGVASORNKNEQNAT 349
DB 297 -----VRS--YMDMEITNTDGTPIIDPTEDTPGICSPDFOGILFGVASORNKNEQNAT 349
QY 301 IDTAPRFGYHWHIQLDNLNGTTPYDPAEDIPGILGTPDFKGI VFGVASQRN--SDGTT 357
DB 301 IDTAPRFGYHWHIQLDNLNGTTPYDPAEDIPGILGTPDFKGI VFGVASQRN--SDGTT 357
QY 350 RAHEAIINTGGDHLCPQISSEIYLT-----PNILRCTNPQPLPQSLG---R 394
DB 350 RAHEAIINTGGDHLCPQISSEIYLT-----PNILRCTNPQPLPQSLG---R 394
QY 358 RAHEAKVDITSGRTFKLGSLEISTESDDFDN-----OPTKFTPVGVDNEADFOQ 417
DB 358 RAHEAKVDITSGRTFKLGSLEISTESDDFDN-----OPTKFTPVGVDNEADFOQ 417
QY 395 TILIRSDNGCHDMVGTSPPTPTWPOQWRRRCSCGNSCCSGHRYPPVPMNVRTWVLVS 454
DB 395 TILIRSDNGCHDMVGTSPPTPTWPOQWRRRCSCGNSCCSGHRYPPVPMNVRTWVLVS 454
QY 418 GPI-----THNMNLAPAVAPNPPGBQLLFFRSQRPSSGGRSNGILDCLVPQEWVQHIF 469
DB 418 GPI-----THNMNLAPAVAPNPPGBQLLFFRSQRPSSGGRSNGILDCLVPQEWVQHIF 469
```


Query Match 58.0%; Score 1679.5; DB 12; Length 548;
Best Local Similarity 59.3%; Pred. No. 7.8e-138;
Matches 336; Conservative 67; Mismatches 119; Indels 45; Gaps 8;

QY 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNNFVQAP 60
DB 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNNFVQAP 60

QY 61 AGEFTVSPRNSPGEILLDELGPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120
DB 61 AGEFTVSPRNSPGEILLDELGPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120

QY 121 FAAIPPGPPYENLSQITMCPHVLIIDVROLEPFLPMPDIWNFFHYNOGNDPKRLVA 180
DB 121 FAAIPPGPPYENLSQITMCPHVLIIDVROLEPFLPMPDIWNFFHYNOGNDPKRLVA 180

QY 181 MLYTPLRANSGDDVFTVSCRVLTKPSPDFEFTFLVPPVTSKTKQFALPILKISEMTNS 240
DB 181 MLYTPLRANSGDDVFTVSCRVLTKPSPDFEFTFLVPPVTSKTKQFALPILKISEMTNS 240

QY 241 RPPVPDVMTARNENQVQNGRVTLDGELLGTTPLLANVICKFKEV---IAKNGD- 296
DB 241 RPPVPDVMTARNENQVQNGRVTLDGELLGTTPLLANVICKFKEV---IAKNGD- 296

QY 297 -----VRSY--RMDMEITNTDGTPIIDPTEDTPGPIGSPDFQILFGVASORKNEONPAT 349
DB 297 -----VRSY--RMDMEITNTDGTPIIDPTEDTPGPIGSPDFQILFGVASORKNEONPAT 349

QY 301 IDTATPRSGFNHWHIQLDNLNGTVPDPAEDIPGLGTPDFKGI VEGVASORN---SDGTT 357
DB 301 IDTATPRSGFNHWHIQLDNLNGTVPDPAEDIPGLGTPDFKGI VEGVASORN---SDGTT 357

QY 350 RAHEAIINTGDDHLCPOISSSEIYLTS-----PNILRCTNPQPLPQSGL---R 394
DB 350 RAHEAIINTGDDHLCPOISSSEIYLTS-----PNILRCTNPQPLPQSGL---R 394

QY 395 GTILIRSDNGHCHDMVGTSPPTTPWQWRRCRSGNSCCSGHRYPPVPMNRVTWIVLS 454
DB 395 GTILIRSDNGHCHDMVGTSPPTTPWQWRRCRSGNSCCSGHRYPPVPMNRVTWIVLS 454

QY 455 HKSFGSTSTRKLPQLNLRLWPLIRFINDTGRVLFPEARLHKOGFITVAHTGNDPIVMPNG 514
DB 455 HKSFGSTSTRKLPQLNLRLWPLIRFINDTGRVLFPEARLHKOGFITVAHTGNDPIVMPNG 514

QY 515 YFRFEAWNOFYSLAPVGTGKRRRVQ 541
DB 515 YFRFEAWNOFYSLAPVGTGKRRRVQ 541

QY 522 YFRFESWNPFTYLLAPMGTGNGRRRIQ 548
DB 522 YFRFESWNPFTYLLAPMGTGNGRRRIQ 548

RESULT 15
Q80FK8 PRELIMINARY; PRT; 548 AA.

AC Q80FK8;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/2197-00/SWE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=227519;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/2197-00;
RA Nilsson M., Hedlund K.-O., Thorhagen M., Johansen K., Ekspong A.,
RA Svensson L.;
RT "Evolution of human calicivirus RNA in vivo: accumulation of mutations
RT in the protruding P2 domain of the capsid leads to structural changes
RT and possibly a new phenotype";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY247435; AAP03033.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 59798 MW; B904034B9F4FEC24 CRC64;

Query Match 58.0%; Score 1679.5; DB 12; Length 548;

Best Local Similarity 59.3%; Pred. No. 7.8e-138;
Matches 336; Conservative 66; Mismatches 120; Indels 45; Gaps 8;

QY 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNNFVQAP 60
DB 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNNFVQAP 60

QY 61 AGEFTVSPRNSPGEILLDELGPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120
DB 61 AGEFTVSPRNSPGEILLDELGPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120

QY 121 FAAIPPGPPYENLSQITMCPHVLIIDVROLEPFLPMPDIWNFFHYNOGNDPKRLVA 180
DB 121 FAAIPPGPPYENLSQITMCPHVLIIDVROLEPFLPMPDIWNFFHYNOGNDPKRLVA 180

QY 181 MLYTPLRANSGDDVFTVSCRVLTKPSPDFEFTFLVPPVTSKTKQFALPILKISEMTNS 240
DB 181 MLYTPLRANSGDDVFTVSCRVLTKPSPDFEFTFLVPPVTSKTKQFALPILKISEMTNS 240

QY 241 RPPVPDVMTARNENQVQNGRVTLDGELLGTTPLLANVICKFKEV---IAKNGD- 296
DB 241 RPPVPDVMTARNENQVQNGRVTLDGELLGTTPLLANVICKFKEV---IAKNGD- 296

QY 297 -----VRSY--RMDMEITNTDGTPIIDPTEDTPGPIGSPDFQILFGVASORKNEONPAT 349
DB 297 -----VRSY--RMDMEITNTDGTPIIDPTEDTPGPIGSPDFQILFGVASORKNEONPAT 349

QY 301 IDTATPRSGFNHWHIQLDNLNGTVPDPAEDIPGLGTPDFKGI VEGVASORN---SDGTT 357
DB 301 IDTATPRSGFNHWHIQLDNLNGTVPDPAEDIPGLGTPDFKGI VEGVASORN---SDGTT 357

QY 350 RAHEAIINTGDDHLCPOISSSEIYLTS-----PNILRCTNPQPLPQSGL---R 394
DB 350 RAHEAIINTGDDHLCPOISSSEIYLTS-----PNILRCTNPQPLPQSGL---R 394

QY 395 GTILIRSDNGHCHDMVGTSPPTTPWQWRRCRSGNSCCSGHRYPPVPMNRVTWIVLS 454
DB 395 GTILIRSDNGHCHDMVGTSPPTTPWQWRRCRSGNSCCSGHRYPPVPMNRVTWIVLS 454

QY 455 HKSFGSTSTRKLPQLNLRLWPLIRFINDTGRVLFPEARLHKOGFITVAHTGNDPIVMPNG 514
DB 455 HKSFGSTSTRKLPQLNLRLWPLIRFINDTGRVLFPEARLHKOGFITVAHTGNDPIVMPNG 514

QY 515 YFRFEAWNOFYSLAPVGTGKRRRVQ 541
DB 515 YFRFEAWNOFYSLAPVGTGKRRRVQ 541

QY 522 YFRFESWNPFTYLLAPMGTGNGRRRIQ 548
DB 522 YFRFESWNPFTYLLAPMGTGNGRRRIQ 548

Search completed: June 1, 2004, 13:53:24
Job time : 32.4978 secs

THIS PAGE BLANK (USPTO)